Psychological Bulletin

HARRY HELSON, Editor
University of Texas

CONTENTS

Stimulus Generalization ..... SARNOFF A. MEDNICK AND JONATHAN L. FREEDMAN 169

The Matching of Behavior Rating and Questionnaire Personality Factors ..... WESLEY C. BECKER 201

Temporal Experience ..... MELVIN WALLACH AND ALBERT I. RABIN 213

Postasymptotic Performance Decrements during Massed Reinforcement ..... WILLIAM F. PROKASY 237

Published Bimonthly by the
American Psychological Association

VOL. 57, NO. 3 MAY, 1960

Consulting Editors

R. R. BLAKE
University of Texas

W. R. GARNER
Johns Hopkins University

J. P. GUILFORD
University of Southern California

W. H. HOLTZMAN
University of Texas

D. W. MACKINNON
University of California, Berkeley

L. J. POSTMAN
University of California, Berkeley

S. B. SELLS
Texas Christian University

W. A. WILSON, JR.
University of Colorado

The *Psychological Bulletin* contains evaluative reviews of research literature and articles on research methodology in psychology. This JOURNAL *does not publish reports of original research or original theoretical articles.*

Manuscripts should be sent to the Editor, Harry Helson, Department of Psychology, University of California, Berkeley 4, California.

Preparation of articles for publication. Authors are strongly advised to follow the general directions given in the *Publication Manual of the American Psychological Association* (1957 Revision). Special attention should be given to the section on the preparation of the references (pp. 50–60), since this is a particular source of difficulty in long reviews of research literature. *All copy must be double spaced, including the references.* All manuscripts should be submitted *in duplicate*, one of which should be an original typed copy; author's name should appear only on title page. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail and to check carefully the typing of the final copy.

Reprints. Fifty free reprints are given to contributors of articles and notes.

ARTHUR C. HOFFMAN
Managing Editor

HELEN ORR
Promotion Manager

SARAH WOMACK
Editorial Assistant

Communications—including subscriptions, orders of back issues, and changes of address—should be addressed to the American Psychological Association, 1333 Sixteenth Street N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Annual subscription: $8.00 (Foreign $8.50). Single copies, $1.50.

PUBLISHED BIMONTHLY BY:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Menasha, Wisconsin

and 1333 Sixteenth Street N.W., Washington 6, D.C.

Entered as second class matter at Washington, D.C., and Menasha, Wisconsin. Second-class postage paid at Washington, D.C., and at additional mailing offices. Printed in U.S.A.

Copyright, 1960, by The American Psychological Association, Inc.

VOL. 57, NO. 3 MAY, 1960

Psychological Bulletin

STIMULUS GENERALIZATION[1]

SARNOFF A. MEDNICK
University of Michigan
AND
JONATHAN L. FREEDMAN
Yale University

Stimulus generalization (SG) is an empirical phenomenon which has, of late, been seeing heavy duty as an explanatory construct in many disparate situations. It has been used in theoretical explanations of discrimination learning (Spence, 1936), of transposition (Spence, 1937), verbal learning (Gibson, 1940), psychoanalytic displacement (Miller, 1948), the behavior of brain-damaged individuals and schizophrenics (Mednick, 1955, 1958a), cross-cultural research (Hull, 1950a), projective techniques (Moylan, 1959), and psychotherapy (Magaret, 1950). It seems likely that these explicatory uses of SG are only a beginning. Inasmuch as we are probably never exposed to exactly the same stimulus situation more than once, all responses to stimulus "repetition" might justifiably be attributed to SG. This "intellectual imperialism" is condoned and encouraged by suggestions that differential reaction time responses (Gibson, 1939) and most of psychophysics (Brown, Bilodeau, & Baron, 1951) may in large part be understood in terms of SG. While we have some sympathy for these viewpoints, the purpose of this article is to provide a concise summary of empirical findings concerning research on SG and to discuss evidence bearing on some important issues.

Organization of the article. After first defining SG we shall review the area of sensory generalization. Here we will be referring to generalization along stimulus continua that can be measured on a physical scale (e.g., distance, intensity, size). Next we will discuss the effect of variables such as drive and degree of training on SG. We shall conclude with a discussion of issues around which much research has revolved.

Definition. Stimulus generalization can be said to occur when a response, previously trained to be elicited by Stimulus 0, can also be elicited by test stimuli similar to 0. The gradient of stimulus generalization (GSG) can be said to be observed

[1] This paper was made possible by grants from the National Science Foundation (G-3855) and the American Philosophical Society (Grant No. 2132) to the senior author. Freedman was at Harvard University when this paper was written; Mednick was at the Institute of Personality Assessment and Research, University of California, Berkeley. The authors wish to thank Tallulah Brown Maki for the important questions and criticisms she raised as a member of a seminar on stimulus generalization conducted by the senior author at Harvard University. Judson S. Brown, Norman Guttman, and Leo J. Postman gave the writer the benefit of their criticism of a preliminary draft of this paper. Roger Shepard's comments were of great help in preparing the mathematical theories section. Penny Reveley contributed invaluable editorial assistance.

if the strength of these generalized responses (measured by frequency, latency, etc.) varies as an orderly function of the physical difference between the test stimuli and Stimulus 0. (This is what we will mean when the terms SG and GSG are used in the remainder of this article.) Through loose usage, these terms have been invested with a rather large number of different connotations. We are in agreement with Brown, et al. (1951) who have pointed up the importance of distinguishing "between the *empirical phenomenon of generalization* on the one hand and *theories or hypotheses about generalization on the other*." The use of the terms in a hypothetical manner connotes non-observable instigative factors while the use of the terms in an empirical manner connotes nothing more than SG and GSG as operationally defined above. Unless the term SG is specifically otherwise qualified, we shall be referring to the empirical usage.

Sensory Generalization

In the main, studies in this area use both the classical and the instrumental conditioning paradigms. After initial conditioning, tests for generalization are made by presenting stimuli which vary in their similarity to the conditioned stimulus (CS). Usually SG testing is continued over a series of extinction trials with some reinforced (booster) trials with the CS interspersed to maintain response strength. Following the custom of the Russian and American literature, the original CS will be denoted Stimulus 0 and generalization stimuli (GS) as Stimulus 1, Stimulus 2, Stimulus 3, etc., in terms of their ordinal similarity to Stimulus 0.

Spatial Generalization

Studies by Bass and Hull (1934), and by Grant and Dittmer (1940), made use of a tactual-vibratory conditioned stimulus applied directly to some point on *S*'s body (Stimulus 0). Shock to the right wrist was used as the unconditioned stimulus (UCS); the galvanic skin response (GSR) served as the conditioned response (CR) and generalization response measure. Tests for generalization were made by applying the tactual-vibratory stimulus to points on *S*'s body at varying distances from Stimulus 0. Thus, instead of varying the disparity between the CS and the UCS with changes in the physical properties of the stimulus, they varied the point of application to the body. A typical situation in the Bass and Hull study had *S*'s left shoulder as Stimulus Point 0, while three other points 16 inches apart on the body were used as generalization test points (small of back, left thigh and left calf). Seventy-four per cent of the trials during generalization testing were reinforced booster trials with Stimulus 0. The other 26% were SG test trials with Stimuli 1, 2 and 3. Measurable GSRs were noted upon application of Stimuli 1, 2 and 3; the amplitude of responses to these stimuli decreased as a function of the distance of the stimuli from Stimulus 0. Thus, SG and GSG as defined above were observed. In a second part of the study, *E* trained all the stimuli (0, 1, 2, 3) to an equal level, extinguished the response to Stimulus 0 to one-fourth of this level and tested 1, 2, and 3 for the generalized effects of this extinction. A gradient of generalization of the effects of extinction was found.

Grant and Dittmer's study paralleled that of Bass and Hull, except that only the back (shoulder to waist) and the dorsum of the right hand were stimulated and the UCS was administered to the left wrist of the group that had the GSR measured on their right hand. They

spaced their four tactual-vibratory stimulus points four inches apart on the back and one inch apart from the wrist to the index finger. The gradients under these conditions were concave upward.

In an attempt to test the applicability of SG findings of conditioning experiments to verbal learning situations, Gibson (1939) carried out a study somewhat similar to the Bass and Hull investigation. Two groups were used, one having the tactual-vibratory stimuli placed 4 in. apart across their backs and the other having them placed 4 in. apart down their backs, forming two separate generalization continua. *S*s were instructed to indicate verbally when the tactual-vibrator was applied to Stimulus Point 0 and to inhibit verbal responses when stimulated at the other stimulus positions. The reaction time of the verbal response was measured by means of a voice key. For both the vertical and horizontal continua the GSGs of frequency of responses showed significant differences between Stimulus Points 0 and 3, 1 and 3, and 0 and 1. No significant gradient of reaction times was obtained. In fact, the average latency of response to the GSs was less than that to Stimulus 0. Gibson explained this as probably being due to the instructions inhibiting long-latency "false" responses to the GSs. Another interesting finding was an upturn at the end of the GSG when vibrators were distributed horizontally; in the horizontal position, Points 0 and 3 were bilaterally symmetrical and this symmetry of location is offered as an "explanation" of the upturn. This explanation was also used by Anrep (1923) when his dogs showed similar responses to bilateral symmetrical stimulation; Anrep concluded that symmetry was an effective elicitor of generalization responses and that stimulation of the point bilaterally symmetrical with Stimulus 0 was as effective as stimulation of the 0 point itself.

It might be maintained that this voluntary verbal response situation is too different from conditioning to be considered an instance of SG and that all Gibson observed was a failure to discriminate Stimulus 0 from the GSs. However, before training, she tested *S*'s ability to discriminate between the stimulus points. The group that discriminated better was the one which later generalized more, casting doubt on any "failure of discrimination" hypothesis. It could also be argued that the so-called generalization responses were nothing more than chance errors which were due to hurrying the response. This argument might have some weight if it were found that speed of reaction and number of generalized responses were positively related. However, there was no such relationship. This same finding for a voluntary response SG study is reported elsewhere (Mednick, 1955). In fact, the group that had slightly longer reaction times was the group that generalized more. In any case, Gibson's operations fulfill our initial definition of SG.

In an attempt to demonstrate spatial generalization in a sensory modality other than tactual, Brown, Bilodeau and Baron (1951) devised a procedure for testing for SG along a visual-spatial continuum. Their apparatus consists of a horizontal row of seven lamps and a reaction time key for *S*. *S* is told that this is a test of reaction time and that when the center lamp, Stimulus 0, is flashed, he should release the reaction key as quickly as he can. He is warned that the other lamps will also be flashed and that he should not respond to these peripheral lamps. After a series of training trials with Stimulus 0, the testing period begins. This period consists of test trials with

the peripheral lamps interspersed among booster trials with Stimulus 0. In two separate experiments symmetrical gradients were found. No tests for significance were made. Here again, however, the latency data showed no consistent trends. Brown, et al. concur with Gibson's interpretation of the lack of a latency gradient when the generalization of a voluntary response is being tested.

This method has been used as a measure of SG in a number of studies. Its use as a measure of SG may be questioned on several grounds. First a study by Slater-Hammel (1955) using a similar apparatus has shown that reaction speed to peripheral visual stimulation decreases as a function of its peripherality. Similar results are reported in Woodworth (1938). Because of the inhibitory instructions used with this method, relatively long reaction times might possibly be expected to be related to reduced responsiveness. (It should be made clear that typically such a relationship has not been found.) Thus it is conceivable that some part of the observed gradient could be attributed to this factor. A second problem concerns the fact that instructions rather than reinforcements produce the major behavioral variations. In view of this, it cannot be said that experiments using this apparatus and method make precise tests of Hull's theory. This is a fact and should be carefully noted, but it does not rule out consideration of the method as one producing "generalization-like effects." These "generalization-like effects" (the term commonly used by Brown and his associates) have proven to be surprisingly parallel to the effects resulting from classical conditioning-generalization studies.

Andreas (1954) pointed out that the empirical gradients reported by Brown, et al. under conditions of inhibitory instructions might be due not only to gradients of associative strength, but could also be ascribed to gradients of inhibitory tendencies (here he is talking in nonempirical "instigative" terms). He tested to see if gradients could be obtained using this type of apparatus without the inhibitory instructions. While the gradients he reports are somewhat irregular, they do show a statistically significant decrement suggesting that positive gradients underlie the observed gradients. Unfortunately, reaction times were not reported so that gradients of latency in the "non-inhibitory instructions" situation could not be observed.

Brown, Clarke and Stein (1958) also criticized the method because the "instructions must be phrased to condone implicitly but not encourage false reactions." They have devised a new procedure using the same basic apparatus which eliminates this problem. The lamps are identified as horses and each trial as a race. *S*'s task is to guess whether the horse (lighted lamp) will win or lose. During training all the lamps but the center lamp win with equal frequency (20%). The center lamp wins more frequently than the others (80%). During the test trials *S*s show a gradient-like pattern of the "win" response with a peak at the center lamp and decreasing frequency of guesses as a function of distance from the center lamp.

Despite the objections, several studies have been reported which have made use of the procedure with the inhibitory instructions. Mednick (1955) repeated the Brown, Bilodeau, and Baron study, extending the stimulus continuum by adding two lamps at each end of the stimulus board for a total of 10 generalization

lamps. He reports a regular, statistically significant gradient of frequency of response. No regular latency gradient was obtained though latency measures were taken.

The failure of studies involving the generalization of voluntary responses to find latency gradients has been of some interest to researchers in this field. It has probably attracted such attention because of the relative ease with which latencies of voluntary responses may be measured (as compared with latencies of GSR responses). Gibson's explanation of this failure has been generally accepted. As briefly pointed out above, she suggested that with inhibitory instructions, potential long latency responses would be withheld if a GS were the test stimulus, but would be allowed to occur to the CS. On the basis of this, Rosenbaum (1953) suggested that the usual gradient might be obtained with latency data if *Ss* were given a set to respond very quickly, thus counteracting the effect of the inhibitory instructions. Mednick (1958c) attempted to establish such a set in his instructions to *S* and found that by only using the data of the 30 *Ss* with the fastest reaction times, he obtained a relatively consistent positively accelerated gradient of latency. In contrast, the gradient of the 30 slowest *Ss* was extremely inconsistent. This study was then replicated on a separate group and the findings were consistent with the hypothesis that the faster *Ss* would demonstrate a "typical" latency gradient of generalization. In a study of the GSG as a function of age in children, Mednick and Lehtinen (1957), using this apparatus, predicted and found that younger children were less restrained by the inhibitory instructions and demonstrated a latency gradient. These two studies seem to support Gibson's interpretation of the failure to find latency GSGs. Other studies using the Brown, Bilodeau, and Baron apparatus, or some modification of it, have investigated the effect of level of manifest anxiety and experimental naivete on SG (Mednick, 1957) summation of GSGs (Bilodeau, Brown & Meryman, 1956), the central tendency effect on GSG (Gewirtz, Jones, & Waerneryd, 1956), the correlation of SG with ethnocentric attitudes (Arnhoff, 1956, 1957), and intelligence (Arnhoff & Loy, 1957). Most of these studies will be discussed below under appropriate headings.

Conclusions: From these studies we might reasonably assert the following with respect to SG along a spatial continuum:

1. With human *Ss*, GSGs can be obtained for both voluntary and involuntary response dimensions with statistically significant gradients regularly observed.

2. In voluntary response situations, with inhibitory instructions, using latency as the response measure, GSGs have not been regularly observed. However, some recent studies have supported explanations of this failure and have reported gradients.

Pitch Generalization

In studies of pitch generalization, some special problems present themselves. If frequency is varied (keeping intensity constant) apparent loudness varies concomitantly. Thus, stimulus intensity dynamism may become a complicating factor. Secondly, human *Ss* have had extensive experience in differentiating pitch so that simple use of cps to characterize points on a pitch similarity continuum seems naive. There are points on the pitch similarity continuum which are well separated in

terms of cps, but are actually difficult for *Ss* to discriminate, e.g., octaves.

Hovland (1937a) handled these problems by establishing a jnd scale using four points separated by 25 jnds as his test stimuli. He varied intensity in order to equate for apparent loudness. In this study (CR was GSR; UCS was electric shock), a regular GSG was observed with significant differences between test points. In a repetition of Hovland's study, Littman (1949) used an improved GSR apparatus. In the early test trials, generalization was complete (responses to test stimuli as great as responses to Stimulus 0) with a significant GSG only appearing in the fourth cycle of SG testing. Littman explains the relative irregularity and extreme height of his plotted gradients as possibly being due to his *Ss* being less acclimated and less experienced in the GSR-shock situation than were Hovland's *Ss*. In support of Littman's interpretation, Mednick (1957) has since shown that generalization responsiveness is indeed greater for relatively naive *Ss*.

In view of the complications involved in testing in more than one octave, Humphreys (1939) questioned the use of as broad a range of test stimuli (153–1967 cps) as employed in the Hovland studies. He restricted his test stimuli to 1000–1967 cps (only 25 jnds). Under 100% reinforcement conditions (such as were used by Hovland and Littman), he noted a concave upward gradient. Those *Ss* receiving 50% reinforcement produced a GSG that was quite elevated. He also found that the octave of Stimulus 0 had an unexpectedly strong response eliciting power (octave effect). A study by Wickens, Schroder, and Snide (1954) also made use of the Hovland paradigm except that only the three lowest tones were employed (153, 468 and 1000 cps) and each *S* was tested on only one GS. Regular GSGs were observed. Le Ny (1957a, 1957b) employed the same tones as Hovland but required a voluntary key-pressing response. He found regular gradients.

Garmezy (1952) used two sets of five tones, one ranging from 500 to 560 cps and the other ranging from 700 to 760 cps. *Ss* were trained either on 515 or 715 cps and were tested with all the GSs in that set. The response consisted of pushing or pulling a handle. Conditions of reinforcement were varied and smooth regular gradients were produced by all conditions of training.

The octave effect was also reported by Blackwell and Schlosberg (1943) who studied rats running for food and away from shock in a straight alley maze. *Ss* were trained to run to all stimuli (10, 8, 7, 5, and 3 kc.) and extinguished on all but the 10 kc. stimulus. A gradient of latency of response was produced with a reversal at 5 kc., the octave of 10 kc. A second study aimed at minimizing the secondary harmonic effects at the octave produced an even more apparent octave effect.

In a study of the effect of various attitudes upon SG, Razran (1949a) made use of a pitch continuum. With his dental cotton technique, he conditioned salivation to the Tone C and tested with the Tones F, B, e, a, d′, g′, $c\sharp^3$, and $f\sharp^3$. While attitude (cooperative, noncooperative) affected conditioning, it did not influence generalization. In Table 1 (p. 823) Razran presents the percent generalization of each of the attitude groups for each of the generalization pitches. Razran suggests that the data indicate that he failed to observe anything other than a very

rough two step gradient. However, inasmuch as attitude had no effect on generalization, we are able to combine the data for all the attitude groups for each of the generalization stimuli. Thus, we may reprocess the reported data to plot a GSG along a continuum of unusual breadth, almost four octaves. The resultant gradient is very regular and steeply descending.

Jenkins and Harrison (1958) made use of an operant conditioning method for studying SG in pigeons which was suggested by Guttman and Kalish (1956). They conducted two experiments in pitch generalization on the effect of differential training (CR extinguished to absence of CS) on SG. They found that differential training produced steeper gradients than nondifferential training.

The following may be reasonably asserted regarding SG along a pitch continuum:

1. Regular and reliable gradients are observed.

2. Augmented responsiveness to the octave of Stimulus 0 has been noted.

3. As with some other continua, initial generalization test trials often elicit augmented SG responsiveness. This is often observed in studies utilizing a noxious UCS. It would be desirable for such studies to institute a pseudoconditioning control to see what role sensitivization is playing in this situation.

4. Humphrey suggests that in view of the complications inherent in working with pitch SG (particularly the octave effect), future research work should consider such continua as hue. However, in view of the quirks of the $\Delta\lambda/\lambda$ incremental hue discrimination curve, nonmonotonic gradients are not to be unexpected in studies of this continuum also.

Intensity Generalization

In a theoretical article, Hull (1949) discusses an aspect of intensity generalization studies which should be initially clarified; other things remaining constant, an organism will react with greater amplitude, frequency, and speed to a stimulus of greater intensity. Since all studies of intensity generalization use training stimuli of different intensity from the generalization test stimuli, according to Hull's postulate, tests for generalization should show strength of reactivity to be a joint function of SG and stimulus intensity. Hull predicts different forms of observed gradients for test stimulus intensities ranging above and below the CS. He predicts convex upward gradients when test stimulus intensities range above the CS, and concave upward gradients when test intensities range below the CS. His postulate also leads to the prediction that the general trend of the gradient originating at a relatively strong stimulus intensity and generalizing toward weaker stimulus intensities will have a greater downward slope than will the gradient extending in the opposite direction between the same stimulus intensities.

The fact that the shape of a curve may be modified by scale transformation mitigates the importance of Hull's notions unless there is some specification of the stimulus and response units. This question is discussed again in the section on the shape of the gradient.

Intensity of direct light source. Hull finds confirmation for his postulate of the joint action of SG and stimulus intensity in a study by Brown (1942). This study may be considered with a group of experiments in which *S*s were trained to respond to a certain intensity of direct light (as opposed to intensity of reflected light) and

tested with other intensities. Brown's study tested the adient pull of rats to the stimulus of a screen of 0.2, 5, or 5000 apparent foot-candles illumination which, in turn, was associated with food. In accordance with Hull's prediction, the GSG (of a maze-running, harness-pulling response) was steeper for "high intensity trained" animals (trained on a high intensity stimulus and tested for SG on lower intensities, CS being the highest stimulus) than for "low intensity trained" animals (trained on the lowest intensity stimulus and tested on higher intensities, CS being the darkest stimulus). These terms, "high intensity trained" and "low intensity trained" will be used henceforth to denote groups trained respectively at the upper or lower extremity of the stimulus continuum used in the specific study.

Fink and Patton (1953) in part of a larger study, also found greater slope of GSG for high intensity trained rats than for those low intensity trained, using rate of water drinking as his response measure with two degrees of direct light (56 foot-candles apart) as the stimulus dimension. Frick (1948), using only high intensity trained rats in generalization test trials with varying degrees of illumination (20 to .02 foot-candles), found a concave upward gradient.

In an experiment with preschool children, Spiker (1956b) differentially reinforced a white (positive) and a blue (negative) light for a lever-pressing response. He then tested for generalization with three additional white lights; half of the *Ss* were high intensity trained and half were low intensity trained. The gradient for the high intensity trained was, in accordance with Hull's prediction, significantly steeper than the gradient for low intensity trained.

These studies, varying intensity of a direct light source, generally support Hull's principle of stimulus intensity dynamism.

Intensity of reflected light source. Another group of studies yielding results somewhat incompatible with those reported above, used intensity of reflected light as the generalization continuum. Raben (1949) trained rats to run in the brightest of five painted runways (varying in steps of 0.33 log units) and then tested for GSG of response latency in subsequent test sessions. She noted that the GSG went from convex upward in the first session (contrary to Hull's prediction) to concave upward in the fourth: this change paralleled a progressive decrement in SG responsiveness. Grandine and Harlow (1948) trained monkeys to get food from under a block of a certain brightness and tested for generalization with blocks of varying brightness by the simultaneous comparison method. They found a significantly decreasing GSG on the first trial; the authors state that the obtained GSGs did not differ significantly from a straight line.

Montgomery (1953), hypothesizing an exploratory drive which can be satisfied by exposure to a novel stimulus situation, extinguished exploratory behavior of rats in one maze and then tested for generalization of extinction to other mazes which varied in the intensity of the light reflected from their surfaces. He obtained an essentially straight line gradient by plotting percentage response decrement against difference of maze illumination. Schlosberg and Solomon (1943) trained rats to establish a black–white dis-

crimination and then tested generalization to five degrees of relative brightness. They found a straight line gradient.

In an ingenious adaptation of the Brown, Clarke, and Stein (1958) method, Bass (1958) used variations in the grayness of the projected silhouette of a jockey on a running horse as her SG continuum. Every exposure of the silhouette represented another race; *S* was required to guess whether the horse would win. The Stimulus 0 horse won 80% of its races, while the other three horses (Stimuli 1, 2 and 3) won 40% of their races. It was assumed and found that the tendency to guess "win" would generalize from Stimulus 0 to the GSs as a function of their distance from Stimulus 0 on the grayness scale. This method could be readily adapted for use with children.

In general, studies using the intensity of direct light source as a continuum tend to support Hull's stimulus intensity dynamism principle; studies using reflected light tend to show ambiguous or negative evidence. Inspection of the two groups of studies suggests that the variations in reflected light (in the range of intensities thus far employed) have not produced as compelling a difference in intensity as variations in the intensity of a direct light source.

Intensity of sound. Hovland (1937b) tested for generalization of a CR to varying degrees of intensity of a tone; CS frequency was 1000 cps, the GSs were each 50 jnds apart and placed 40, 60, 74, and 86 dbs. above the threshold. Human *Ss* were used with PGR as CR and electric shock as UCS. Hovland trained half of the *Ss* to low intensity, half to high, with a counterbalanced design. While he does not plot differential gradients for low intensity trained as opposed to high intensity trained, it is apparent from inspection of the reported data that the slope of the gradient for the high intensity trained group is greater than that for low intensity trained. Fink and Patton (1953), studying learned drinking responses in white rats, found that the GSG for high intensity trained *Ss* tended to have steeper slope than that for low intensity trained *Ss*.

Miller and Green (1954) tested for generalization of extinction effects along a continuum of buzzer intensity. Using rats motivated by shock to run to a buzzer signal along a T maze, they reported a GSG with greater slope for high intensity trained animals. The gradient of extinction trials for the low intensity trained group approached a horizontal line, which result is explained in part by the greater number of trials which this group took to learn as compared with the high group.

From studies of sound intensities, the theory of SG responsiveness as a joint function of SG and stimulus intensity is supported.

Drive intensity SG. Stimulated by some studies showing indirect evidence of the generalization of responses to proprioceptive stimulation associated with specific durations of food deprivation, Yamaguchi (1952) explicitly designed a factorial study to test for this phenomenon. He deprived rats of food for from 3 to 72 hrs. during training and tested for generalization along the continuum of proprioceptive stimulation resulting from different degrees of hunger at each deprivation level. He found significant GSGs; the slope of the GSG of the high intensity trained group was sig-

nificantly greater than the slope of the GSG for those that were low intensity trained. What is of special interest is the fact that generalization could be measured along this unusual continuum.

Some new interpretations. Johnsgard (1957), following up some work by Bragiel and Perkins (1954), has presented evidence suggesting that it is the degree of stimulus-background contrast rather than absolute stimulus intensity which is the variable producing the stimulus intensity dynamism effect. Johnsgard had rats run on an elevated runway to stimulus cards of varying reflectance placed before a screen of relatively medium intensity. He made more accurate predictions of number of responses and of latency from consideration of the contrast of stimulus card and background than from the absolute intensity of the stimulus card.

Perkins (1953) has also suggested a new interpretation (with supporting evidence) of the observed effects of stimulus intensity. He maintains that during the conditioning procedure an inhibitory response is built up to the absence of the CS, i.e., response is inhibited at those times that CS intensity equals zero. The inhibitory GSG originating at the point of zero intensity dampens response strength to the relatively weaker intensities of the CS. He further points out, in agreement with Heyman (1957), that in the typical maze running, direct light-intensity study, another factor to be considered is the fact that *S* is experiencing the highest possible intensity of the stimulus at the moment of reward. Since the animal usually gets the reward near the light source, it makes sense that he should respond with greater vigor to higher stimulus intensities. This interpretation may also explain the ambiguity of the findings in the reflected light intensity studies.

Hull (1950b, 1951, 1952) himself, has presented a revision of his original ideas on stimulus intensity dynamism. This revision has received little research attention. The revision suggests that intensity enters as a variable both during training (affecting ${}_{s}H_{R}$) and during response evocation. (In the earlier paper no effect on ${}_{s}H_{R}$ was postulated.) This revision actually has only a minor effect on his earlier predictions. Heyman (1957) in an excellently conceived study, evaluated the revised position. He used high intensity and low intensity trained rats approaching and pushing open a door covered with grey stimulus paper of appropriate reflectance. He found regular GSGs with slope and shape in agreement with Hull's predictions. He did not find that habit strength (${}_{s}H_{R}$) was affected by the intensity of the training stimulus. His results suggest that Hull's earlier position is better able to predict his results than the revised model.

Conclusions. The following conclusions may be reasonably asserted regarding SG along an intensity continuum:

1. Statistically significant gradients are regularly observed.

2. The slope of the GSG for *S*s trained at high intensity and tested with GSs of low intensities is steeper than that for *S*s trained with the CS at low intensity and tested at high intensity.

3. It is possible that studies using reflected light as the stimulus continuum do not have enough effective intensity differences, in comparison with direct light studies, to yield

results in accordance with Conclusion 2.

4. Some recent work suggests that stimulus background contrast is an important variable determining the effective intensity of the stimulus.

Size Generalization

A number of experiments have investigated generalization along the dimension of stimulus size. The basic paradigm is similar to that of other SG studies; *S* is trained to respond to a CS of a given size and is then tested for generalization to stimuli above and/or below the CS on the particular size dimension being employed. According to most theoretical formulations, the amount of generalizations should be inversely proportional to the distance of the given stimulus from the CS. Although this expectation is, to a large extent, supported by the data, there are various factors which complicate the situation and which make it difficult to obtain regular gradients.

The two earliest studies were conducted by Gulliksen and by McKinney. Gulliksen (1932) used size and direction of angles as the stimuli and the moving of a handle to the left or right as the response. A gradient along the size dimension was reported. McKinney (1933) employed geometric line drawings that varied in the total length of the lines used. The results show "that the percentage of transfer decreases as the amount of alteration increases" (p. 857).

More recently, a series of experiments has been conducted employing circles of varying areas as stimuli. All of these studies used an apparatus designed by Grice, consisting of a straight alley maze 24 in. long with a starting box at one end and the stimulus at the other end. The stimulus circle has a small door through which a rat may reach the feeding dish. The maze can be rotated so that *S* does not have to be replaced at the end of a trial.

In the initial study by Grice and Saltz (1950), five white circles (79, 63, 50, 32, and 20 sq. in.) were used. *Ss* were trained either to the largest or smallest circle, and were tested on all stimuli in an extinction situation. Bidirectional GSGs were obtained.

Kling (1952) investigated generalization of extinction using the Grice apparatus. *Ss* were trained to respond to two different stimuli which were presented singly. One of the two stimuli was then presented alone for 15 nonreinforced trials. Testing consisted of a presentation of the other stimulus. The measure of generalization was the latency of the first response to the test stimulus. Smooth gradients were obtained for generalization from large to small and from small to large.

Margolius (1955) repeated Grice's experiment varying the number of reinforced trials during the training periods, and employing only the 79 sq. in. stimulus as the CS. Using the mean latency of the first three test trials as a measure, smooth gradients were obtained. Two other measures —number of responses in 30 trials and number of responses in 60 sec.— also produced regular gradients which weren't as precise as the latency gradient.

A study by Brush, Bush, Jenkins and Whiting (1952) is quite similar to the three just discussed. With pigeons as *Ss*, an illuminated spot of varying size on a pecking target in a Skinner box was used as the stimulus. Generalization was tested on both sides of the CS; rate of response was

the response measure. Asymmetrical gradients were produced, that on the larger side being higher. However, if the log of size was substituted for absolute size, the gradients become more symmetrical. Jenkins, Pascal, and Walker (1958) also used this operant technique and report smooth and fairly regular GSGs.

Quite a few studies have investigated SG using the height of the stimulus as the independent variable. Grandine and Harlow (1948) trained monkeys to find food in a dish under a block of a given height. The *S*s were then presented with the CS and a stimulus of a different height. The choice of the nonrewarded stimulus was considered an SG response. First trial data yielded regular GSGs.

A study by Buss (1950) followed one of the basic concept formation paradigms. *S*s were presented with blocks of various heights, colors, areas, etc., and were made to respond "vec" or "non vec" to each block. Two response keys corresponding to the two possible responses were provided. During a short training period all "vec" responses to tall blocks were reinforced. Generalization gradients were produced along the height dimension both by number and latency of the "vec" responses. This is one of the few experiments in this area, or in fact in the whole area of stimulus generalization, in which human *S*s, making a voluntary response, produced a gradient based on latency. It is interesting to note that there were two positive responses (i.e., lifting either key), rather than a single positive response and inhibiting instructions as in comparable studies. Two other studies (Buss, Weiner & Buss, 1954; Buss, 1955) making use of the same basic method and varying mode of reinforcement also produced smooth GSGs.

Rosenbaum (1953) using psychiatric and normal *S*s, employed rectangles of light of various heights as the stimuli. The response consisted of moving a slider along a wire. Strong and weak shock as well as a buzzer were used to increase the speed of the response and to heighten anxiety. Regular gradients of amplitude were reported for all conditions except for the group of normal *S*s that received the buzzer. A nonsignificant reverse gradient of latency was found, i.e., latency varied inversely as a function of GS dissimilarity. Unfortunately, during the testing session three CS booster trials were regularly presented between test stimuli. This would tend to prepare *S* for the test stimulus and constitutes a factor whose influence in the experiment is difficult to assess.

In a study by Botwinick (1953), *S*s were trained to a small rectangle and were tested for generalization to three larger rectangles. The response consisted of moving a handle to the right. The measures used were frequency, amplitude and latency. For the first trial, first 10 trials and for the first 20 trials gradients were reported for frequency and amplitude. Using all trials, a gradient was found for latency, as well as for the other measures.

Grant and Schiller conducted a study with the GSR as the response measure, and shock as the UCS. A rectangle of light 1 in. $\times$ 12 in. was the CS and rectangles of 9, 10, 11, 13, 14, and 15 in. were the test stimuli. A gradient was reported based on first trial data, but a reversal occurred on the largest stimulus (1953).

In general, regular gradients have been observed for generalization

along the size dimension. In some cases, however, reversals and asymmetries have been observed. These have uniformly occurred in cases where *Ss* have been trained to a relatively small CS and tested on larger GSs. It seems likely, as pointed out by Grice and Saltz, that this is due to a general effect which tends to augment responsiveness to larger stimuli. In view of the fact that large stimuli appear brighter than small stimuli, the effect may be indirectly due to stimulus intensity dynamism.

An alternative or complementary explanation could be based on the problem of defining equal distances between stimuli. If equal physical distances are used, it is evident that the psychological distances between the smaller stimuli are greater than those between the larger stimuli. This would make the larger stimuli closer to each other and to the CS than are the smaller stimuli, producing an elevated gradient. Brush et al. (1952), have demonstrated that using logarithmic scales tends to compensate for this effect.

Conclusions. From these studies, we can reasonably assert the following with respect to SG along the dimension of size:

1. For both human and infrahuman *Ss*, for both voluntary and involuntary response measures, for measures based on frequency, amplitude and in some instances latency of response, regular GSGs are observed.

2. The use of a slider response for human *Ss*, such as used by Gulliksen, Rosenbaum, and Botwinick, provides excellent amplitude and frequency data. However, latency GSGs are not observed.

3. A group of studies from widely separate laboratories which made use of the Grice apparatus have obtained highly consonant and meaningful results. This recommends this method quite highly for animal work.

Drive Stimulus Generalization

In this area we ask: (*a*) whether a response can be trained to the proprioceptive stimulation arising from a deprivation state; and (*b*) if so, will generalization of this response occur to proprioceptive stimulation which arises from similar deprivation states. In an attempt to answer these questions, Yamaguchi (1952) used a modified factorial design, Skinner box bar pressing response, and varying degrees of internal response to food deprivation as stimuli, with rats as *Ss*. He found generalization along this drive stimulus continuum.

Temporal Generalization

Pavlov (1927) demonstrated that dogs could learn to respond to a time interval as a conditioned stimulus. Several studies have subsequently investigated SG along this dimension. Czehura trained four dogs to respond to a one second silent interval between two tones differing in pitch. He then tested responsiveness to three additional durations of silent interval—3, 9, and 26 sec.—observing a regular gradient.

Rosenbaum (1951) trained rats to press a bar every 60 sec. to obtain food pellets. During testing, seven additional intervals were used—14, 30, 45, 75, 90, 105, and 120 sec. Using a latency measure of response strength, a significant gradient, decreasing in height as the CS was approached, was obtained; past the 60 sec., Stimulus 0 point, latencies tended to remain at the same level as those for the CS interval. The

lack of a regular gradient beyond the CS interval was interpreted as the result of *S* being set to respond at 60 sec., at which point internal stimulus cues would be maximal. Therefore, for intervals greater than 60 sec. response latencies would not be expected to increase. A question which can be raised regarding this study is the positioning of the animal. As the testing interval approached the CS interval, the animal could have been moving closer to the bar, so that proximity would result in decreased latency. After 60 sec. the animal might have remained poised at the bar resulting in the observed plateau.

A further study of the temporal dimension was that by Fink and Davis (1951) with human *Ss* giving a voluntary key-pressing response to three series of tones of differing duration. Muscle action potential data was also recorded. In one series, *Ss* were told to respond to all tones and were presented with tones of 2, 2.25, 2.5, 3.0 and 4.0 sec. In the second series (both given to all *Ss*) only two-second tones were presented. During testing, 10 trials were given with the two-second durations, then 10 trials with the other four durations. Fairly regular convex upward gradients of both muscle action potential and key pressing were obtained. It is of interest in this study that the key pressing response yielded results which were directly comparable to the results obtained from the muscle action potential response. These results on muscle action potential in a sense support the use of voluntary response measures with human *Ss*. Fink and Davis discuss the problem of choosing an interval scale that is most appropriate for the independent variable. They suggest that any measure is as good as another if followed consistently; logs are of particular value for human *Ss* when the temporal intervals used are very small.

A final study in this area is that by Mednick (1959), who used taped tones differing only in duration (CS = 2.7 sec., GSs = 3, 3.3, 3.6, 3.9 sec.) Twenty CS training trials were given and a regular generalization gradient of a voluntary key pressing response was found during testing.

It may be concluded that regular, statistically significant GSGs may be obtained along a temporal dimension. It would seem to be advantageous for *Es* to adopt a consistent method of presenting gradients in terms of the interval scale expressing the independent variable. Log units appear to be the most generally applicable.

Hue Generalization

Utilizing a modified version of the Skinner automatic key-pecking pigeon apparatus, Guttman and Kalish, are conducting a series of careful investigations of SG along the hue continuum. The stimulus is produced by a monochromator and displayed on the translucent pecking key. Typically, the pigeon is brought to a high level of responsiveness by means of a variable interval reinforcement schedule and generalization testing is carried on under extinction conditions. Regular gradients are almost invariably obtained with this technique. It also has the advantage of enabling *E* to obtain GSGs from individual *Ss*.

In an early experiment Guttman and Kalish (1956) studied the relationship between the degree of discriminability of changes in a stimulus

continuum and the shape of the generalization gradient along that continuum. One might predict an inverse relationship between the incremental discriminability of a continuum (as measured by $\Delta\lambda$) and the amount of generalization. However, these variables were found to be relatively independent of each other. Kalish (1958) adapted this apparatus for humans and found an inverse relationship between generalization and discriminability along the wave length continuum. The differences in procedure and *Ss* between the Guttman and Kalish (1956) and Kalish (1958) studies makes it difficult to explain these opposing results.

In a later study on summation of gradients (Kalish & Guttman, 1957) the method proved successful in producing regular GSGs along the hue continuum. An outstanding doctoral dissertation by Honig (1958), using the same basic procedure investigated problems in transposition and stimulus preference. Of special interest is the finding that stimulus preference may be predicted from knowledge of the underlying generalization gradients. Hanson (1957) had pigeons learn to peck at a key for one wavelength and not to peck for another. The pigeons were then tested for generalization along the wave length continuum that surrounded these two stimuli. As might be expected, generalization responsiveness on the side of the negative stimulus was greatly reduced. What is surprising is that the discrimination training did *not* reduce the total amount of generalization responsiveness. Instead the entire GSG seemed to move up the continuum away from the negative trained stimulus. The maximum point of the curve was also displaced in this direction. In an informative paper on the problem of the scaling of the independent variable, Guttman (1956) reports on some current work from his laboratory.

Variables Influencing Generalization

Stimulus Generalization as a Function of Drive Level

Studies in this area have concerned themselves with the effect of hunger, sex, and fear-anxiety on the GSG.

Hunger. Brown (1942), using the straight alley maze, rat-pulling-harness apparatus, varied the number of hours of pretest food deprivation. He noted regular gradients with the amplitude and speed of response varying directly with number of hours of food deprivation. The GSG of the high drive group also tended to be steeper. In a later study of similar design (Brown, 1948), these findings were replicated. An additional condition of strong electric shock produced results which suggest that the gradient of avoidance (shock) behavior is steeper than that for approach (hunger) behavior.

Rosenbaum (1951), studying generalization along a temporal dimension, using a modified Skinner box, varied drive by depriving two groups of rats for 22 hours and then feeding one group a small quantity of food just before testing. He noted no significant differences between the GSGs of his two groups when the fed group received only 4 gms. of food. When this amount was increased to 10 gms. in a second experiment, the difference between the groups was significant. The unfed group demonstrated greater SG responsiveness.

The more pigeons are deprived of food the more generalization re-

sponsiveness they exhibit (Jenkins, Pascal, & Walker, 1958). The pigeons were tested using operant techniques with the size of the pecking key serving as the generalization continuum.

None of these studies utilized a factorial design; this factor would tend to complicate the interpretation of their findings due to the possibility of a drive stimulus SG effect. Despite this, their results make it reasonable to conclude that increased time of food deprivation will lead to increased SG responsiveness.

Electric shock. Rosenbaum (1953) used three degrees of noxious stimulation (strong shock, weak shock, and buzzer) to produce varying levels of drive state with human *Ss*. While the weak shock and buzzer conditions did not produce differing results, the strong shock resulted in considerably elevated gradients.

In one of the series of studies on displacement, Murray and Miller (1952) compared the GSG under three levels of electric shock, finding greater generalization responsiveness for the greater shock condition. Brown (1948) reported a similar finding. Relatively strong shock produced relatively elevated gradients.

Sex. Testing the effects of testosterone proprionate upon the copulatory behavior of sexually inexperienced male rats exposed to "incentives" of varying degrees of "appropriateness," Beach (1942) measured responsiveness on a 5 point scale. A rough gradient may be observed from his data. The greater the dose of testosterone proprionate the less the incentive need resemble a receptive female in order for it to elicit a response and the stronger the response to the incentive. Macerone and Walton (1938) obtained similar results.

Situational factor. Mednick (1957) compared the GSGs of experimentally naive and experimentally sophisticated *Ss* finding that the naive *Ss* demonstrated greater SG responsiveness than the sophisticated *Ss*. Inasmuch as this effect was especially marked in *Ss* who scored high on the Taylor Manifest Anxiety Scale he interpreted this finding in terms of situational drive arousal.

It may be reasonably concluded from the above evidence that any condition which will increase the drive state will result in increased SG responsiveness.

Individual Differences in Generalization

The study of individual differences is concerned with a search for correlates of SG reactiveness. The bulk of the work has dealt with developmental factors, personality variables, schizophrenia, and brain damage. In most instances the intent is utilization of SG as an explanatory concept. Thus an excess in SG reactiveness is seen by some writers as an important basis of the disorder of thinking in schizophrenia.

However, there is another possible direction that this research might take. Basic problems in learning could be attacked by this method. If a theory hypothesized SG as an important variable promoting learning in a given context, then a group evidencing a high level of SG reactiveness would be expected to perform well in such a context. For example, Gibson (1940) hypothesized that SG is behind the interference arising in the learning of serial lists of high intralist similarity. A group showing reduced SG reactiveness

should learn a high intralist similarity list with little interference (relative to its performance on a low intralist similarity list). Mednick (1955) has shown that individuals with cortical brain damage evidence reduced SG reactiveness. On the basis of this result, Carson (1958) hypothesized and demonstrated that cortically brain damaged individuals showed relatively little decrement in learning as a function of increasing intralist similarity. In studies of this type, it is necessary that the effect be shown in the interaction term since simple differences between groups could be due to a myriad of other factors.

Manifest anxiety. Mednick (1957), comparing high and low scoring college students on the Taylor Manifest Anxiety Scale (MAS) found that the highly anxious *Ss* showed more spatial generalization. Studies by Hilgard, Jones, and Kaplan (1951) on eyelid conditioning, Wenar (1953) on temporal conditioning, and M. T. Mednick (1957) on PGR conditioning and mediated generalization, have also demonstrated greater generalization reactiveness on the part of high anxious *Ss*. In the study by M. T. Mednick it was found that conditioning level of the anxiety groups was directly related to degree of anxiety. In order to evaluate generalization responsiveness independent of conditioning level, she compared anxiety groups equated for conditioning level; the nature of her findings was not changed. It is important that all studies attempting to demonstrate group differences in SG responsiveness either equate their groups for conditioning level or use a measure of relative generalization (ratio of generalization response to conditioned response).

A study by Buss (1955), on psychiatric patients found no differences as a function of MAS scores; he suggests that the MAS is not an appropriate test for psychiatric patients. Fager and Knopf (1958) agree with Buss' conclusion on the basis of their experiment with psychiatric patients. They found no relationship between MAS scores and SG.

Clinical groups. On the basis of clinical observation, Schilder (1939) first remarked on the difficulties schizophrenics have in tasks involving differentiation. Bender and Schilder (1930), studying conditioned withdrawal from shock, noted extreme over-generalization by schizophrenic *Ss*. Cameron (1951) has spoken of over-inclusion and the broadening of the GSG on the part of schizophrenics. Garmezy (1952), studying generalization along the dimension of pitch, found that schizophrenics demonstrated more relative generalization than normals. This effect was especially marked under conditions of socially administered punishment. Mednick (1955), using the dimension of space, found that schizophrenics generalized more than normals. Dunn (1954) tested schizophrenics with social and nonsocial materials and found relatively greater generalization for schizophrenics with the social materials. The "social" generalization continua were very cleverly constructed, in one instance, consisting of changes in the angle of elevation of the pointing arm of a scolding mother. Rosenbaum (1953) utilized psychiatric patients in a generalization study, but did not compare them with the normal *Ss* in the study. However, inspection of the reported data suggests that there was little apparent difference in generaliza-

tion between the groups. Mednick (1958a) has presented a learning theory approach to research in schizophrenia which leans heavily on SG theory.

Eriksen (1954) compared size generalization in "hysterics" and "psychasthenics" as defined by extreme scores on the appropriate MMPI scales. The "hysterics" demonstrated more generalization. Both groups show more generalization when instructed that they could avoid the UCS (electric shock) than when told it was unavoidable.

Personality tests. Arnhoff (1956) reported that high generalizers showed more ethnic prejudice (E scale), but on replication of his study (1957) he found the difference did not hold up. In this laboratory, similar fluctuating results have been observed across a group of four studies using the highly related F scale (authoritarianism). Analysis of the subscales of the F scale indicates that the Projectivity subscale consistently holds up. High "projectors" tend significantly, to be high generalizers. In the process of carrying out generalization research at this laboratory, various additional personality measures have been incidentally correlated with generalization responsiveness. We will mention these briefly. No relationship has been noted between SG responsiveness and field independence (Witkin, 1954) or need achievement (McClelland, Atkinson, Clarke, & Lowell, 1953); on the other hand, those low in achievement imagery (Iowa Picture Interpretation Test, Hedlund, 1953), tend to generalize more than those high in achievement imagery (Mednick, 1958b).

Age. Mednick and Lehtinen (1957) found that SG responsiveness varied inversely as a function of age in school children; no relationship was found for adults in a separate study by Mednick (1955). With the younger children, Mednick and Lehtinen noted a latency gradient while none was noted for the older children. In line with these results, Reiss (1946) found generalization along a gradient of homonymy to vary inversely as a function of age in children.

Brain damage. In a study of patients suffering from a variety of types and locations of cortical brain damage, Mednick (1955) noted extreme curtailing of generalization responsiveness. These results are similar to those of Smith (1951) who noted decreased transfer (based on deficiency in SG responsiveness) from *Ss* with corpus callosum damage. Goldstein's (1941) concept of the "concreteness" of the brain damaged may also be based on some retardation in SG responsiveness. A study by Mednick and Wild (1958) on cerebral palsied children indicates an accompanying deficiency in SG responsiveness.

Degree and Type of Training

SG has been regarded as a spread of habit strength from the CS to the GSs. Thus, one of the variables that might influence the GSG is level of habit strength of the CR. Experiments in this area have manipulated either amount or distribution of reinforcement of the CR.

Number of reinforced trials. The earliest relevant data comes from Pavlov's laboratory and is reported by Razran (1949b). Using dogs as *Ss*, the salivary CR, and an assortment of stimuli as CS, Pavlov found, in a number of different experiments, that the amount of generalization to test stimuli increased with increased

number of training trials to the CS.

Beritoff (1924), experimenting with dogs, reported a "spread of generalization," that is, as training continued, tones further and further removed from the CS could elicit the response. However, after three days of training generalization ceased to spread. There was considerable variation among *S*s; no statistical analysis was reported.

In a study by Hovland (1937c), different groups were given 8, 16, 24, or 48 reinforcements. Generalization occurred at all levels of training; relative generalization (i.e. ratio of generalized response to conditioned response strength) increased with training and then decreased. *S*s receiving 16 reinforcements generalized not only more than those receiving less reinforcement, but also more than those receiving more reinforcement. It was also found that the more reinforcement *S*s received, the more rapidly the generalized response extinguished. Hovland (1935) conducted a similar experiment, the results of which again indicate that relative generalization increases from 8 to 16 reinforcements and then decreases.

Razran (1940) experimented with the conditioned salivary response in humans using the increase in weight of cotton dental pads placed in *S*'s mouth as the response measure. Conditioning consisted of short periods during which *S*s ate while the appropriate CS was presented. In the tests for generalization, which were held at the end of each of the eight training sessions, different colored lights, buzzers and monosyllabic words were substituted for the CS. There was generalization to each stimulus at each degree of training. Relative generalization increased with training and then declined; the decline was more marked when single stimuli were used as CS than when combinations of stimuli were used. The interpretation of the results of these experiments is somewhat difficult. For one thing, while Razran states that there were from 2 to 16 *S*s in each study, 22 of the 26 studies had only 2 *S*s. Further, when summarizing the various results he takes the mean of a group of percentages based on differing numbers of *S*s.

In a study by Spiker (1956a) children were trained to push a lever in order to get marbles which they could exchange for toys. Pushing the lever was rewarded with marbles only when a particular color of light was presented in contiguity with the response; each correct response was rewarded. Two groups of *S*s, equated for response frequency, were given either 12 or 24 training trials. At the end of the training period generalization was tested with different colored lights. The group that received more reinforcement responded to the generalization test stimulus more often.

The experiments discussed so far dealt with the effect of amount of reinforcement on the amount of generalization, the latter being tested with only a single GS. An excellent study by Margolius (1955) investigated the relationship between number of reinforcements and the gradient of generalization. Using the Grice apparatus, *S*s (white rats) were trained on a 79 sq. cm. circle and tested on circles of 63, 50, 32, and 20 sq. cm. Groups of *S*s were given different numbers of reinforcements ranging from 4 to 104. In testing for generalization, three measures were used: number of responses in 30 trials, latency of the first three test trials, and number of responses in 60

sec. Absolute and relative generalization increased with training, the smoothest gradients being obtained from the latency measure.

In the studies reported, SG responsiveness increased as a function of the number of reinforced training trials. On the other hand, those studies using human *Ss* and PGR or salivary CR measures found relative generalization increasing up to a point and then decreasing. Margolius studied rats running in the Grice maze apparatus and reported a direct relationship between relative generalization responsiveness and number of reinforced training trials. These differences in results are ascribable to too many differences in procedure and sampling to allow for unequivocal interpretation. However, there is a possible interpretation. The conditioned PGR is highly sensitive to the effects of adaptation, far more so than reinforced maze running on the part of hungry rats. After 24 or 48 reinforced trials, a gross overall adaptation of the PGR would not be unexpected. This could easily have masked any possible advantage that the generalized response was gaining from increased training.

Distribution of reinforcement. In a study by Humphreys (1939), the percentage of reinforcement of the CR was varied. With tones as stimuli, PGR as CR, *Ss* received either 100% or 50% reinforcement. All *Ss* received the same number of trials. The 100% reinforced group produced a negatively accelerated gradient; the 50% reinforced group responded almost equally to all stimuli, including the CS, though there was a slight decrease in responsiveness to the stimulus farthest removed from the CS. In a subsequent repetition of this study which made use of a voluntary response Humphreys found essentially the same results (1948). Humphreys concluded that it is reasonable to predict a positively accelerated GSG for the 50% reinforced group when the gradient becomes more marked. This group also generalized more than the 100% reinforced group, despite the fact that the latter received twice as many reinforcements. In explanation, Humphreys points out that the 50% group was concerned about the shock during the generalization trials, while the other group was not. Another possible explanation of the results is that the extinction situation was more similar to the training situation for the partially reinforced group than it was for the 100% reinforced group (i.e. from 50% reinforcements to no reinforcements is less of a change than from 100% reinforcements to no reinforcements).

Theories of Stimulus Generalization

Neurophysiological Formulations

While intuitively it would seem that SG is a phenomenon which affords fertile ground for explanation couched in terms of neurophysiology, little attention has been devoted to such attempts. Pavlov (1927) first put forth the notion that SG is due to a fading wave of excitation irradiating from the spot on the cortex stimulated because of the presentation of the CS. As a result of this wave of excitation, stimuli whose cortical representations are contiguous to that of the CS also possess the potentiality of eliciting the CR. Due to the fading strength of the wave, the response eliciting power of these stimuli decreases as a function of

their distance from the cortical representation of the CS.

Loucks (1933) and Denny-Brown (1932) have developed analytical arguments which seem to cast doubt on the "irradiation" hypothesis. Even more damaging, however, is the data from a study by Grant and Dittmer (1940). They point out that the cortical representation of the hand is considerably greater in area than that of the trunk of the body. From this, they infer that irradiation will suffer a greater decrement between two comparable points on the hand than between two points on the trunk. Their data showed the gradient for the trunk to be slightly steeper. Grant and Dittmer also point to the Bass and Hull study as tending to refute the irradiation hypothesis. Bass and Hull (1934) report a smooth gradient going from shoulder to feet. In terms of irradiation theory this is extremely surprising since the cortical representations of the thigh, calf and foot are interposed between the cortical representation of the waist and buttocks. Some aberrancy in the GSG should be observed as a result of this. While these data and analytic arguments cast strong doubt on the usefulness of the irradiation hypothesis as it stands, it certainly is appealing for heuristic reasons and may prove important in some modified form.

Wolpe (1952) has attacked the problem from the standpoint of excitation of common neural pathways. He gives an example from experiments on pitch generalization. It is known that tones which are similar to each other in terms of cps will tend to excite overlapping parts of the relevant end organ of hearing. Further, this end organ overlap is mirrored at higher neural levels such as the medial geniculate body and the cerebral cortex. Thus, a response learned to one of these tones may be called forth by the other tone by means of these common stimulated pathways. Tone A might stimulate Nerve Endings 1, 2, and 3 with a peak at 2. Tone B will stimulate 3, 4, and 5, with a peak at 4. Any response called forth by Tone A will also be elicited by Tone B because of Common Pathway 3. A problem which quickly arises in this formulation is the usual finding of octave generalization, i.e., incremented response to the octave of the CS despite the fact that the tones are well separated on the cps dimension. Wolpe explains this as being the result of overtones having the frequency of the CS being produced by presentation of its octave. This explanation suffices for this aspect of the Humphreys (1939) experiment (which Wolpe cites by way of illustration) but does not explain the Blackwell and Schlosberg (1943) finding of a strong-octave effect in rats despite extensive precautions taken to minimize the appearance of such overtones. Wolpe would also have some difficulty explaining increments of generalization response to symmetrical stimuli (Gibson, 1939; Anrep, 1923). Further, the Humphreys experiment, cited by Wolpe, itself contains data damaging to his theory. Humphreys' experiment fully replicated an earlier experiment by Hovland except for one important detail. Hovland tested across the pitch dimension using a range of 4 tones from 1967 to 153 cps. Humphreys only tested between 1967 and 1000 cps. Since the representation of 1967 and 153 cps are more widely separated on the hearing end organ than 1967 and 1000 cps, we are clearly in

a position to predict that in the instance that 1967 was the CS, the 1000 cps test tone of the Humphreys study should elicit more relative generalization than the 153 cps tone of the Hovland study. As it happens the relative generalization strength of Humphreys' 1000 cps tone was .72; the relative generalization strength of Hovlands' 153 cps tone was .70. Humphreys suggests that this similarity of response strengths might be ascribed to a "frame of reference" effect since the tones were both end points of the pitch dimension in their respective studies. These data and such an explanation are not compatible with the Wolpe position.

The common pathways notion is similarly damaged by consideration of the dimension of intensity. Since degree of intensity of a stimulus is not reflected in the firing of a specific neurone it is not clear how common pathways can develop in the intensity dimension.

Form of the GSG

There have been a number of attempts to specify the form of the GSG. These attempts have often made the explicit or implicit assumption that a form exists which is invariant across dimensions. In addition, many have maintained, with Hovland, that "the form of the gradient is of considerable importance in psychological theory" (1937b). Certainly such work as Spence's analysis of transposition, Miller's displacement theory and work on summation of gradients could proceed from firmer foundations if the form or forms of the GSG were better delineated.

Unfortunately there has not been good agreement regarding the shape. In fact, just about all reasonably imaginable shapes have been proposed as "the" form of the gradient (including no constant form at all).

While this question of shape has consumed paragraphs in many a discussion of experimental results, it has only occasionally assumed the status of a central problem. This is perhaps as it should be for the true shape may well be an ephemeral treasure. One dismaying notion is the ease with which any obtained shape may be altered by simply manipulating the units of the axes. It seems clear that without some specification of stimulus and response measurement scales, discussion of the shape of obtained gradients must proceed very cautiously. This fact has been pointed out by Guttman in an interesting treatment of this and other SG problems (1956).

An approach to the problem of the shape of the GSG in terms of a mathematical model is presented by Shepard (1958a) and discussed below.

Mathematical Formulation of SG

The application of mathematical models to SG has been impeded by the absence of a satisfactory independent variable, i.e., a mathematically precise measure of psychological similarity of stimuli. Recently, however, there have been two kinds of attacks on this problem. One of these, proposed by Bush and Mosteller (1951) is based upon a set-theoretic model, and views a stimulus as a collection of elements. The similarity of two stimuli is defined in terms of the fraction of these elements that the two stimulus-sets have in common.

Following Estes (1950) they maintain that the probability of a response to a given stimulus is equal to the proportion of elements in that stimulus conditioned to that re-

sponse. The probability that any new stimulus will elicit this response is proportional to the similarity of the new stimulus to the conditioned stimulus. An index of similarity is determined by application of a set-theory equation to previously obtained data. By introducing further assumptions about the effects of reinforcement and extinction on the primitive elements, set-theoretic models of this kind have been shown to account for certain ways in which generalized response probabilities change during the course of discrimination learning (Green, 1958; Restle, 1955). However, Bush and Mosteller conclude that their model cannot make any experimentally testable prediction about the shape of the GSG since the index of similarity (that is presumed to govern SG) is "very much organism determined" and, hence, has no invariant relation to the distance between stimuli along "such physical dimensions as light or sound intensity, frequency, etc." Restle and Beecroft (1955) have pointed up the similarity "between the Hull-Spence theory and the Bush-Mosteller mathematical model in the study of stimulus generalization" as it is affected by differences in anxiety level.

Shepard (1957) has proposed a model for SG that regards a stimulus as a point in metric space. The similarity of two stimuli is defined in terms of the distance between these stimulus points in their common space. He supposes that stimuli differing along a single physical dimension (e.g. tones of various frequencies) can be identified with points in a "psychological space" in such a way that the distance between points is invariantly related to SG. Such a curve in psychological space differs from the rectilinear physical scale by some "continuous, differentiable transformation." Therefore, although the GSG is postulated to have an invariant form in this psychological space, when this space is retransformed back into the "physical space" (e.g., into the physical scale of frequency) the form of the GSG becomes irregular and depends upon the position of the CS.

He then asks this question: Given the physical scale with its empirically determined but irregular gradients of SG, is it possible to recover the psychological space and, hence, a unique GSG having the same form for all positions of the CS? Shepard's contention is that for any given experimental procedure the GSG has a unique form that is independent of the position of the CS and of the physical dimension along which the test stimuli are arrayed.

His procedure attempts to discover the unique function that will convert the conditional response probabilities from experiments on SG into psychological distances. On the basis of several paired-associate experiments in which the stimuli vary along such physical dimensions as size, brightness, color and shape, he concludes that the GSG under this kind of experimental condition is concave upward (Shepard, 1958a, 1958b). In addition, though, he concludes that the shape of the GSG depends upon the distribution of reinforcement so that, if the S is reinforced only part of the time, the GSG departs from the pure exponential form and becomes more "bell-shaped," i.e., convex upward in the vicinity of the CS. This conclusion is supported by Humphreys' results (1939, 1948). Although this model accounts for the observed form of the GSG (when that gradient is plotted against psychological distance), it

does not attempt an account of the changes in the degree of SG during discrimination learning. In this respect, then, the metric model is less general than the set-theoretic models, like Bush and Mosteller's which incorporate a learning mechanism. Perhaps the two types of approach will eventually be combined so as to account both for the shape of the GSG and for the temporal course of SG.

Summation of Generalized Response Strength

A theorem of Hull's (1939) which has excited much interest but little research suggests that overlapping GSGs originating from two or more points on the same continuum will summate and result in incremented generalization responsiveness in the area of overlap. Bilodeau, Brown and Meryman (1956) studied generalization along a visual-spatial dimension with a task in which *S*s gave an instructed voluntary response. They found a summative effect between two positive points separated by 16° of visual angle but not between two points separated by 32° of visual angle. They are careful to point out that while their procedures have a formal correspondence to the methods assumed by Hull, they do not meet all of his assumptions. Guttman and Kalish (1957) point out that the Bilodeau, Brown and Meryman experimental design permits summation to be observed at six different points on the gradients. It occurs at only one of these, the single point between the training stimuli. This latter finding is repeated in the data presented by Guttman and Kalish. They used the pigeon pecking apparatus exposing different hues on the pecking key to test for generalization. Their method comes closer to satisfying the conditions for summation as assumed by Hull (e.g., reinforcement used to build up habit strength rather than instructions) than that of Bilodeau, Brown and Meryman. Thus, while in both studies summation seems observable at points between the training stimuli, it is disturbing not to find it at any of the other possible points on the gradient.

The Lashley, Wade-Hull Controversy and Some New Interpretations

Hull (1943) postulated that as a result of the reinforcement of a response to a stimulus, gradients of associative strength develop to nonconditioned stimuli which produce empirical GSGs. In an uneven article which, however, contains penetrating criticisms demanding response, Lashley and Wade (1946) present a strong attack on this Hullian postulate as well as other aspects of research on SG. One point lies at the basis of their criticism; this is their assertion that the empirical phenomena of SG do not mirror underlying gradients of habit strength developed during conditioning, but instead represent a "failure of association." By this they mean that the reason *S* responds to the GS in the test period is that *S* has not yet been conditioned to respond differentially to the *E*-defined relevant aspect of the CS. It is only during testing that the GSG develops, and this is only because of *S*'s attention being directed to the *E*-defined dimension.

Hull's (1947) otherwise effective and informative answer to the Lashley and Wade paper failed to deal with the important "failure of association" argument. Brown, Bilodeau, and Baron (1951) fractionate the argument and dismiss it as circular

and untestable. They point out that the only way one can know when *S* has failed to learn to respond to a particular stimulus element is when *S* does not respond differentially to that characteristic. However, this failure to respond differentially is the operational definition of SG. Since these two are defined by identical operations, Brown, et al. argue that the Lashley-Wade criticism is circular. However, if this criticism were to suggest a specific differential test, it could avoid this property of circularity.

The "failure of association" notion is in one sense similar to Estes' learning theory since it proposes that only certain elements of the stimulus situation are associated with the response on any one learning trial, due to the specificity of *S*'s attending responses. If, in a test trial, *E* changes only one small aspect of the stimulus, *S* will respond to the aspect to which he had attended and which remains unchanged. However, if, during training, *S* has by chance attended to the *E*-defined relevant element of the stimulus then *S* will not respond when that element is changed; no SG will be observed. In these terms a specific prediction may be made. The greater the number of conditioning trials the greater the likelihood of *S* attending to the relevant stimulus element and thus the greater the likelihood of differentiation of the CS from the GS. Thus, generalization would be expected to decrease as a function of the number of conditioning trials. As pointed out above quite the opposite result has been regularly observed. This is seen as very telling evidence against the failure of association position.

Growing from the failure of association argument is the assertion that the GSG is a product of the test period and develops as a result of *S*'s attention being directed to the *E*-defined relevant dimension. Lashley and Wade assert that no falling GSG will be observed unless the CS is contrasted with a GS so that the relevant dimension is made manifest to *S*. Since underlying gradients of response strength cannot develop in the conditioning period "training with one stimulus fails to produce a significantly greater strength of association with that stimulus than with others on the same dimension." Lashley and Wade specifically suggest the method of single stimulus training (followed by a single test with a GS) as a crucial test. Such a test has since been repeated many times over, resulting in occasional reports of first test trial gradients. The Grandine and Harlow (1948) study was specifically designed with the Lashley and Wade restrictions in mind. Regular GSGs were observed on the first test trial. However, in this instance it should be noted that Grandine and Harlow are in the minority. Most experiments have not found a falling gradient on the first test trial. More often there is no difference between the CS and the GSs.

Razran (1949) has proposed another interpretation of generalization phenomena. He suggests that there are two types of generalization: "(a) *pseudogeneralization* and (b) *true generalization*." The Lashley-Wade *failure of association* position refers to pseudogeneralization phenomena where *S*, because of age, infirmity, or circumstances, is unable to note distinguishing characteristics of the CS. "The relation of pseudogeneralization to true generalization is not unlike that of the undifferentiated total action of the young foetus to the structured whole activities

of the fully developed individual."

In his "subsequent testing hypothesis" Razran further asserts that "generalization develops, not during the original training of the conditioned stimuli, but during the subsequent testing of the generalization stimuli." This hypothesis is similar to the Lashley-Wade position on this matter. Razran is suggesting that during training no tendencies or propensities to generalization are built up. During testing with the GSs, *S*s will "categorize or rate the new stimulus on some sort of crude similarity-dissimilarity scale" consisting of two or three steps.

The present writers hold a view somewhat intermediate between that of Lashley, Razran and Hull. Before developing this view it may be helpful to analyze some of the assumptions of what we will call, for want of a better name, the Hull position. It should be made clear that this elaboration of Hull's statements is the writers' responsibility. It is not impossible that in order to draw a sharper contrast they have in certain instances exaggerated Hull's position.

Vigorous generalization responses are routinely observed on the first test trial following single stimulus training. Inasmuch as such responses would not occur in the absence of this training and the degree of such training has a marked effect on the degree of generalization behavior, the phenomenon of stimulus generalization is probably best seen as resulting from the operations involved in training.

On the first test trial following training (generally acknowledged to be the "purest" measure of SG) one should expect to find a regular descending GSG. Instead, the GSs usually elicit almost as much response strength as the CS ("complete" SG). Hull suggests that this is due to strong responsiveness occurring immediately after training (not necessarily a convincing argument in view of the action of reactive inhibition which would tend to lower responsiveness) and to unextinguished reactions to irrelevant situational stimuli such as apparatus clicks. Wickens, Schroder and Snide (1954) attempted to extinguish such responses to an irrelevant aspect of the stimulus before beginning SG testing, but nevertheless found complete SG on the first test trial.

The expectation of a descending GSG on the first test trial carries with it some assumptions. First, one must assume that SG responsiveness develops as a consequence of training. This seems to be an acceptable assumption. The second assumption is that training results in a gradient of apportionment of potential responsiveness along the relevant stimulus continuum. For example, it is expected that if *S*s are trained to salivate to a tone 1000 cps and each *S* is presented once with a tone at some other point on the cps continuum, the extent of each *S*'s response will be proportional to the physical or jnd distance of his test stimulus from the training stimulus. If *S* is to respond in a proportional manner (i.e., respond more to 1400 cps than to 1800 cps) it must be assumed that he came to the first test trial with associations between specific stimulus values on a physical or jnd scale and certain levels of response strength. That is, it must be assumed that before the first test trial he already had a certain number of drops of salivation set aside for 1400 cps and a smaller number reserved for 1800 cps.

However, there is another possi-

bility. What if, following training, *S* set up associations between generalization responses and an *ordinal* scale of stimulus values rather than a jnd or continuous physical scale? That is, what if the extent of *S*'s response to a given GS was proportional to the *number of stimulus units* that separate that GS from Stimulus 0 (with number of stimulus units defined in terms of the population of stimulus units *S* has experienced within the immediate experimental situation)? This would explain the fact that almost complete SG is usually observed on the first test trial; the first GS that *S* experiences that is on the same continuum as Stimulus 0 will be, at that point, only one stimulus unit from Stimulus 0, and will elicit a relatively large response. For example, let us say that in an experiment on pitch generalization, Stimulus 0 is 1000 cps and one *S* receives a 1400 cps GS and another *S* receives an 1800 cps GS on the first test trial. For each *S* in terms of his experience in the experimental situation, his GS will be only one unit removed from Stimulus 0. With continued testing *S*'s GS units hierarchy will change until he has experienced the full range of GSs used in the experiment. At this point a regular descending GSG would be expected. This would in turn explain another typical finding of generalization experiments, a regular descending GSG is usually only found at a relatively late stage of testing. Also, it is typically observed that the slope of the GSG increases with continued testing.

Let us take another situation where some differential predictions may be made. The Hull position, as stated above, would tend to predict that the amount of responsiveness elicited by a GS will be determined by its distance from the CS along a physical or jnd scale without regard to what other GSs, if any, are presented in the experiment. In terms of the alternative notion being presented here, if the CS is a tone of 1000 cps and 1600 cps is the critical test GS, it will make a great difference whether the other GSs are (*a*) 1200 cps and 1400 cps or if the other GSs are (*b*) 1800 cps and 2000 cps. In the case of (*a*), the GS at 1600 cps will be three units from the CS; in the case of (*b*), 1600 cps will be one unit from the CS. If *S*s allot their generalization responsiveness according to units of separation of GSs from the CS then it would be predicted that the 1600 cps stimulus would elicit more responsiveness in the (*b*) situation than in the (*a*) situation.

Further, it would be predicted that an important determinant of amount of response to any of the GSs would be their units separation from the CS. Thus, the 1200 cps GS in (*a*) should elicit about as much generalization as the 1600 cps GS in (*b*) since both are one unit of separation from the CS. Some empirical support for this interpretation of generalization phenomena is found in a comparison of studies by Hovland (1937) and Humphreys (1939). Hovland reports the generalization of a PGR response to tones 25, 50, and 75 jnds from the CS. Part of the Humphreys experiment repeated the Hovland procedure, confining itself to GSs that were 5, 15, and 25 jnds from the CS. The CS and 25 jnd tones were 1967 dv and 1000 dv, respectively. The two studies are compared in terms of the units hypothesis in Table 1. The units hypothesis would predict that the Humphreys 25 jnd tone would elicit as much response as the Hovland 75 jnd tone (both being three units from

TABLE 1
GENERALIZATION STIMULI IN JNDS FROM CS

Hovland	CS			25	50	75
Humphreys	CS	5	15	25		

the CS). The generalization responsiveness in the two studies may be roughly compared in terms of relative generalization (response to GS /response to CS). The Humphreys 25 jnd GS/CS ratio was .72; the Hovland 75 jnd GS/CS ratio was .70. The 25 jnd Hovland GS (one unit of separation) elicited more relative generalization than the 25 jnd Humphreys GS (three units of separation). Humphreys termed this the "frame of reference" effect stating that the "generalized response may depend upon the setting in which stimuli are perceived." The writer has recently completed a number of studies testing the units hypothesis in SG in a number of continua, spatial, tactual and temporal and in terms of predictions regarding the summation of gradients (1959). An early study (1956) which replicated the Humphreys-Hovland comparison in the spatial continuum using the Brown, Bilodeau, and Baron (1951) apparatus lends strong support to the units interpretation.

This research on the units hypothesis has shown that this is an important variable in determining generalization behavior. However, it has been noted that the predictions of the units variable tend to break down at the extremes of the continuum being tested. An *S* trained to respond to 50 cps and tested at 20,000 cps will not behave the same as an *S* trained at 50 cps and tested at 100 cps despite the fact that the test stimulus is one unit of separation away in both cases.

REFERENCES

ANDREAS, B. G. Empirical gradients of generalization in a perceptual-motor task. *J. exp. Psychol.*, 1954, **48**, 119–122.

ANREP, G. V. The irradiation of conditioned reflexes. *Proc. Roy. Soc., Ser. B*, 1923, **94**, 404–425.

ARNHOFF, F. N. Ethnocentrism and stimulus generalization. *J. abnorm. soc. Psychol.*, 1956, **53**, 138–139.

ARNHOFF, F. N. Ethnocentrism and stimulus generalization: A replication and further study. *J. abnorm. soc. Psychol.*, 1957, **55**, 393–394.

ARNHOFF, F. N., & LOY, D. L. Relationship between two measures of stimulus generalization: Influence of intelligence upon performance, *Psychol. Rep.*, 1957, **3**, 465–470.

BASS, B. Gradients in response percentages as indices of nonspatial generalization. *J. exp. Psychol.*, 1958, **56**, 278–281.

BASS, M. J., & HULL, C. L. The irradiation of a tactile conditioned reflex in man. *J. comp. Psychol.*, 1934, **17**, 47–65.

BEACH, F. A. Effects of testosterone propionate upon the copulatory behavior of sexually inexperienced male rats. *J. comp. Psychol.*, 1942, **33**, 227–248.

BENDER, L., & SCHILDER, P. Unconditioned reactions to pain in schizophrenia. *Amer. J. Psychiat.*, 1930, **10**, 365–384.

BERITOFF, J. S. On the fundamental nervous processes in the cortex of the cerebral hemispheres. I. The principle stages of the development of the individual reflex: Its generalization and differentiation. *Brain*, 1924, **47**, 109–148.

BILODEAU, F. A., BROWN, J. S., & MERYMAN, J. J. The summation of generalized reactive tendencies. *J. exp. Psychol.*, 1956, **51**, 293–298.

BLACKWELL, H. R., & SCHLOSBERG, H. Octave generalization, pitch discrimination and loudness thresholds in the white rat. *J. exp. Psychol.*, 1943, **33**, 407–419.

BOTWINICK, J. Stimulus generalization of human voluntary responses: The relationship between amount of training and generalization of visual stimuli varied in the size dimensions. *Dissertation Abstr.*, 1953, **13**, 1269–1270

BRAGIEL, R. M., & PERKINS, C. C., JR. Conditioned stimulus intensity and response speed. *J. exp. Psychol.*, 1954, **47**, 437–441.

BROWN, J. S. The generalization of approach responses as a function of stimulus intensity

and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.

BROWN, J. S. Gradients of approach and avoidance responses and their relation to level of motivation. *J. comp. physiol. Psychol.*, 1948, **41**, 450–466.

BROWN, J. S., BILODEAU, E. A., & BARON, M. R. Bidirectional gradients in the strength of a generalized voluntary response to stimuli on a visual spatial dimension. *J. exp. Psychol.*, 1951, **41**, 52–61.

BROWN, J. S., CLARKE, F. R., & STEIN, L. A new technique for studying spatial generalization with voluntary responses. *J. exp. Psychol.*, 1958, **55**, 359–362.

BRUSH, F. R., BUSH, R. R., JENKINS, W. O., & WHITING, J. W. M. Stimulus generalization after extinction and punishment: An experimental study of displacement. *J. abnorm. soc. Psychol.*, 1952, **47**, 633–640.

BUSH, R. R., & MOSTELLER, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, **58**, 413–423.

BUSS, A. H. A study of concept formation as a function of reinforcement and stimulus generalization. *J. exp. Psychol.*, 1950, **40**, 494–503.

BUSS, A. H. Stimulus generalization as a function of clinical anxiety and direction of generalization. *J. abnorm. soc. Psychol.*, 1955, **50**, 271–273.

BUSS, A. H., WEINER, M., & BUSS, E. Stimulus generalization as a function of verbal reinforcement combination. *J. exp. Psychol.*, 1954, **48**, 433–436.

CAMERON, N. Perceptual organization and behavior pathology. In R. R. Blake & G. V. Ramsey (Eds.), *Perception, an approach to personality*. New York: Ronald, 1951.

CARSON, R. C. Intralist similarity and verbal rote learning performance of schizophrenic and cortically damaged patients. *J. abnorm. soc. Psychol*, 1958, **57**, 99–106.

CZEHURA, W. S. The generalization of temporal stimulus patterns on the time continuum. *J. comp. Psychol.*, 1943, **36**, 76–90.

DENNY-BROWN, D. Theoretical deductions from the physiology of the cerebral cortex. *J. Neurol. & Psychopath.*, 1932, **13**, 52–67.

DUNN, W. L. Visual discrimination of schizophrenic subjects as a function of stimulus meaning. *J. Pers.*, 1954, **25**, 48–64.

ERIKSEN, C. W. Some personality correlates of stimulus generalization under stress. *J. abnorm. soc. Psychol.*, 1954, **49**, 561–565.

ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.

FAGER, R. E., & KNOPF, I. J. Relationship of manifest anxiety to stimulus generalization. *J. abnorm. soc. Psychol.*, 1958, **57**, 125–126.

FINK, J. B., & DAVIS, R. C. Generalization of a muscle action potential response to tonal duration. *J. exp. Psychol.*, 1951, **42**, 403–408.

FINK, J. B., & PATTON, R. M. Decrement of a learned drinking response accompanying changes in several stimulus characteristics. *J. comp. physiol. Psychol.*, 1953, **46**, 23–27.

FRICK, F. C. An analysis of an operant discrimination. *J. Psychol.*, 1948, **26**, 93–123.

GARMEZY, N. Stimulus differentiation by schizophrenic and normal subjects under conditions of reward and punishment. *J. Pers.*, 1952, **20**, 253–276.

GEWIRTZ, J. L., JONES, L. V., & WAERNERYD, K. Stimulus units and range of experienced stimuli as determinants of generalization discrimination gradients. *J. exp. Psychol.*, 1956, **52**, 51–57.

GIBSON, E. J. Sensory generalization with voluntary reactions. *J. exp. Psychol.*, 1939, **24**, 237–253.

GIBSON, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.

GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, **53** (2, Whole No. 239).

GRANDINE, L., & HARLOW, H. F. Generalization of the characteristics of a single learned stimulus by monkeys. *J. comp. physiol. Psychol.*, 1948, **41**, 327–339.

GRANT, D. A., & DITTMER, D. G. An experimental investigation of Pavlov's cortical irradiation hypothesis. *J. exp. Psychol.*, 1940, **26**, 299–310.

GRANT, D. A., & SCHILLER, J. J. Generalization of the conditioned galvanic skin response to visual stimuli, *J. exp. Psychol.*, 1953, **46**, 309–313.

GREEN, E. J. A simplified model for stimulus discrimination. *Psychol. Rev.*, 1958, **65**, 56–63.

GRICE, G. R., & SALTZ, E. The generalization of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1950, **40**, 702–708.

GULLIKSEN, H. Transfer of response in human subjects. *J. exp. Psychol.*, 1932, **15**, 496–516.

GUTTMAN, N. The pigeon and the spectrum and other perplexities. *Psychol. Rep.*, 1956, **2**, 449–460.

GUTTMAN, N., & KALISH, H. I. Discriminability and stimulus generalization. *J. exp. Psychol.*, 1956, **51**, 79–88.

HANSON, H. M. Discrimination training effect on stimulus generalization gradient for

spectrum stimuli. *Science*, 1957, **125**, 888–889.

HEDLUND, J. L. Construction and evaluation of an objective test of achievement imagery. Unpublished doctoral dissertation. State Univer. of Iowa, 1953.

HEYMAN, W. Certain relationships between stimulus intensity and stimulus generalization. *J. exp. Psychol.*, 1957, **53**, 239–248.

HILGARD, E. R., JONES, L. V., & KAPLAN, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–100.

HONIG, W. Prediction of preference, transposition, and transposition reversals from stimulus generalization. Unpublished doctoral dissertation, Duke Univer., 1958.

HOVLAND, C. I. The effects of varying amounts of reinforcement upon the generalization of conditioned responses. *Psychol. Bull.*, 1935, **32**, 731–732.

HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148. (a)

HOVLAND, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291. (b)

HOVLAND, C. I. The generalization of conditioned responses: IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 261–276. (c)

HULL, C. L. The problem of stimulus equivalence in behavior theory. *Psychol. Rev.*, 1939, **46**, 9–30.

HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.

HULL, C. L. The problem of primary stimulus generalization. *Psychol. Rev.*, 1947, **54**, 120–134.

HULL, C. L. Stimulus intensity dynamism (V) and stimulus generalization. *Psychol. Rev.*, 1949, **56**, 67–76.

HULL, C. L. A primary social science law. *Sci. Mon.*, 1950, **17**, 221–228. (a)

HULL, C. L. Behavior postulates and corollaries, 1949. *Psyhcol. Rev.*, 1950, **57**, 173–180. (b)

HULL, C. L. *Essentials of behavior*. New Haven: Yale Univer. Press, 1951.

HULL, C. L. *A behavior system: An introduction to behavior theory concerning the individual organism*. New Haven: Yale Univer. Press, 1952.

HUMPHREYS, L. G. Generalization as a function of method of reinforcement. *J. exp. Psychol.*, 1939, **25**, 361–372.

HUMPHREYS, L. G. The generalization of verbal expectations following two conditions of reinforcement. *Amer. Psychologist*, 1948, **3**, 347.

JENKINS, H. M., & HARRISON, R. H. Auditory generalization in the pigeon. Final Report, *Air Force Contr. AF 18(603)-85*, 1958.

JENKINS, W. O., PASCAL, G. R., & WALKER, R. W., JR. Deprivation and generalization. *J. exp. Psychol.*, 1958, **56**, 274–277.

JOHNSGARD, K. W. The role of contrast in stimulus intensity dynamism (V). *J. exp. Psychol.*, 1957, **53**, 173–179.

JOHNSON, D. M. Generalization of a scale of values by the averaging of practice effects. *J. exp. Psychol.*, 1944, **34**, 425–436.

KALISH, H. I. The relationship between discriminability and generalization: A re-evaluation. *J. exp. Psychol.*, 1958, **55**, 637–644.

KALISH, H. I., & GUTTMAN, N. Stimulus generalization after equal training on two stimuli. *J. exp. Psychol.*, 1957, **53**, 139–144.

KLING, J. W. Generalization of extinction of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1952, **44**, 339–346.

LASHLEY, K. S., & WADE, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, **53**, 72–87.

LE NY, J. F. Généralisation d'une attitude dans une epreuve de temps de reaction. *Année psychol.*, 1957, **57**, 11–21. (a)

LE NY, J. F. Généralisation d'une attitude dans une epreuve de temps de reaction. II. Cas d'une reaction discriminative. *Année psychol.*, 1957, **57**, 329–337. (b)

LITTMAN, R. A. Conditioned generalization of the galvanic skin response to tones. *J. exp. Psychol.*, 1949, **39**, 868–882.

LOUCKS, R. B. An appraisal of Pavlov's systematization of behavior from the experimental standpoint. *J. comp. Psychol.*, 1933, **15**, 1–47.

MCCLELLAND, D. C., ATKINSON, J. W., CLARKE, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century 1953.

MACERONE, C., & WALTON, A. Fecundity of male rabbits as determined by "dummy matings." *J. Agric. Sci.*, 1938, **28**, 122–134.

MCKINNEY, F. Quantitative and qualitative essential elements of transfer. *J. exp. Psychol.*, 1933, **16**, 854–864.

MAGARET, A. Generalization in successful therapy. *J. consult. Psychol.*, 1950, **14**, 64–70.

MARGOLIUS, G. Stimulus generalization of an instrumental response as a function of the number of reinforced trials. *J. exp. Psychol.*, 1955, **49**, 105–111.

MEDNICK, M. T. Mediated generalization

and the incubation effect as a function of manifest anxiety. *J. abnorm. soc. Psychol.*, 1957, **55**, 315–321.

MEDNICK, S. A. Distortions in the gradient of stimulus generalization related to cortical brain damage and schizophrenia. *J. abnorm. soc. Psychol.*, 1955, **51**, 536–542.

MEDNICK, S. A. Generalization as a function of manifest anxiety and adaptation to psychological experiments. *J. consult. Psychol.*, 1957, **21**, 491–494.

MEDNICK, S. A. A learning theory approach to research in schizophrenia. *Psychol. Bull.*, 1958, **55**, 316–327. (a)

MEDNICK, S. A. Stimulus generalization as a function of level of achievement imagery. *Psychol. Rep.*, 1958, **4**, 651–654. (b)

MEDNICK, S. A. Latency generalization gradients of a voluntary response. *Amer. J. Psychol.*, 1958, **71**, 752–755. (c)

MEDNICK, S. A. Studies in stimulus generalization. Prog. Rep., NSF-Grant G-3855, 1959.

MEDNICK, S. A., & BRADBURN, N. M. On the definition of "space" in spatial generalization. *Amer. Psychologist*, 1956, **11**, 439.

MEDNICK, S. A., & LEHTINEN, L. F. Stimulus generalization as a function of age in children. *J. exp. Psychol.*, 1957, **53**, 180–183.

MEDNICK, S. A., & WILD, C. Stimulus generalization in brain damaged children. Final Rep., USPHS grant M-1519, 1958.

MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.

MILLER, W. C., & GREENE, J. E. Generalization of an avoidance response to varying intensities of sound. *J. comp. physiol. Psychol.*, 1954, **47**, 136–139.

MONTGOMERY, K. C. Exploratory behavior as a function of "similarity" of stimulus situation. *J. comp. physiol. Psychol.*, 1953, **46**, 129–133.

MOYLAN, J. J. Stimulus generalization in projective test (Rorschach) behavior. *J. Pers.*, 1959, **27**, 18–37.

MURRAY, E. J., & MILLER, N. E. Displacement: Steeper gradient of generalization of avoidance than of approach with age of habit controlled. *J. exp. Psychol.*, 1952, **43**, 222–226.

PAVLOV, I. P. *Conditioned reflexes*. Oxford: Oxford Univer. Press, 1927.

PERKINS, C. C. The relation between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, **46**, 225–231.

RABEN, M. W. The white rat's discrimination of differences in intensity of illumination measured by a running response. *J. comp. physiol. Psychol.*, 1949, **42**, 254–272.

RAZRAN, G. H. S. Studies in configurational conditioning: V. Generalization and transposition. *J. genet. Psychol.*, 1940, **56**, 3–11.

RAZRAN, G. H. S. Attitudinal determinants of conditioning and of generalization of conditioning. *J. exp. Psychol.*, 1949, **39**, 820–829. (a)

RAZRAN, G. H. S. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, **46**, 337–365. (b)

REISS, B. F. Genetic changes in semantic conditioning. *J. exp. Psychol.*, 1946, **36**, 143–152.

RESTLE, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, **62**, 11–19.

RESTLE, F., & BEECROFT, R. S. Anxiety, stimulus generalization, and differential conditioning: A comparison of two theories. *Psychol. Rev.*, 1955, **62**, 433–437.

ROSENBAUM, G. Temporal gradients of response strength with two levels of motivation. *J. exp. Psychol.*, 1951, **41**, 261–267.

ROSENBAUM, G. Stimulus generalization as a function of level of experimentally induced anxiety. *J. exp. Psychol.*, 1953, **45**, 35–43.

SCHILDER, P. The psychology of schizophrenia. *Psychoanal. Rev.*, 1939, **26**, 380–398.

SCHLOSBERG, H., & SOLOMON, R. L. Latency of response in a choice discrimination. *J. exp. Psychol.*, 1943, **33**, 22–39.

SHEPARD, R. N. Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, 1957, **22**, 325–345.

SHEPARD, R. N. Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychol. Rev.*, 1958, **65**, 242–256. (a)

SHEPARD, R. N. Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *J. exp. Psychol.*, 1958, **55**, 509–523. (*b*)

SLATER-HAMMEL, A. T. Reaction time to light stimuli in the peripheral visual field. *Res. Quart.*, 1955, **26**, 82–87.

SMITH, K. V. Learning and associative pathways of the human cerbral cortex. *Science*, 1951, **114**, 117–120.

SPENCE, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.

SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.

SPENCE, K. W. The basis of solution by chim-

panzees of the intermediate size problem. *J. exp. Psychol.*, 1942, **31**, 257–271.

SPIKER, C. C. The effects of number of reinforcements on the strength of a generalized instrumental response. *Child Developm.*, 1956, **27**, 37–44. (a)

SPIKER, C. C. The stimulus generalization gradient as a function of the intensity of stimulus lights. *Child Developm.*, 1956, **27**, 85–98. (b)

WENAR, C. Reaction time as a function of manifest anxiety and stimulus intensity. *J. abnorm. soc. Psychol.*, 1953, **48**, 129–134.

WICKENS, D. D., SCHRODER, H. M., & SNIDE, J. D. Primary stimulus generalization of the GSR under two conditions. *J. exp. Psychol.*, 1954, **47**, 52–56.

WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality through perception.* New York: Harper, 1954.

WOLPE, J. Primary stimulus generalization: A neuro-physiological view. *Psychol. Rev.*, 1952, **59**, 8–10.

WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.

YAMAGUCHI, H. G. Gradients of drive stimulus (S_D) intensity generalization. *J. exp. Psychol.*, 1952, **43**, 298–304.

(Received April 21, 1959)

PSYCHOLOGICAL BULLETIN
VOL. 57, NO. 3, 1960

THE MATCHING OF BEHAVIOR RATING AND QUESTIONNAIRE PERSONALITY FACTORS

WESLEY C. BECKER
University of Illinois

Over the years, Cattell has outlined and followed a strategy for determining and assessing personality traits which initially called for the identification of factors in three media, namely, behavior ratings (BR or L data), questionnaires (Q) and objective tests (T). Factor identification was to be based upon blind rotation to simple structure, which according to Cattell has the virtuous property of revealing functional unities which are invariant within the tolerances set by sampling and measurement errors. Implicit in Cattell's strategy was the hope that the various media would reveal similar structurings and thus give evidence, in terms of converging operations, that a structural model had been found which had a claim to fitting reality better than other possible models.

Assertions made in recent publications by Cattell and his co-workers would lead the reader to believe that research has demonstrated an almost one-to-one correspondence between BR and Q personality factors. These assertions also carry with them the implication that Cattell's long sought for evidence of factor invariance across media has been found, at least for BR and Q factors. In his most recent book (1957) Cattell summarizes the present state of his research on cross-media matching with the following statement:

> The upshot of this cross-media matching is clearly that L and Q factors are far more completely mutually matched than either is with T factors. Only L-data factors J and K, and the factors noted before to be specific to Q-data, namely, Q_1, Q_3, Q_3, and Q_4, fail of cross-matching (p. 325).

Cattell hedges this conclusion by presenting possible alternative cross-matches for 3 of 12 factors for which matching is claimed (Cattell, 1957, Table 8–7, p. 326) and by stating in his preface (p. xi) that the matching across media may be even less reliable than it seems. However, an even stronger statement on this subject appears in a recent paper by Scheier and Cattell (1958). They write: "Factors in the rating and questionnaire realm have been securely linked to each other with all but a few factors now in one realm finding their counterparts in factors in the other realm" (p. 608). Further along in this paper, the following statement is made: "Moreover, the linkage between questionnaire and rating realms is so well established that a match to a questionnaire factor is excellent presumptive evidence of a match to a parallel rating factor" (p. 609).

Since the present writer had quite a different impression of the present status of research on the matching of BR and Q factors from that given in the preceding quotations, he undertook the present review of the evidence and issues involved in relating BR and Q factors.

THE EVIDENCE

Scheier and Cattell (1958), and Cattell, Saunders, and Stice (1957) make reference to Cattell (1957) in support of their assertions quoted above. Cattell (1957) cites three cross-media studies, two with adults and one with children.

The first cross-media study was

carried out in 1945–46 by Cattell and Saunders (1950). The sample consisted of 118 college men and 240 college women. The behavior rating factor scores were based on one to five rating variables each, generally those found common to two previous studies (Cattell, 1947; Cattell, 1948). On the average, three rating variables were used to estimate each of 13 behavior rating factors. Questionnaire factor scores were based on one question for eight factors (Q-1, Q-2, Q-3, Q-17, Q-9, Q-4, Q-13, and Q-14) and on five questions for the remaining eight factors (Q-10, Q-5, Q-8, Q-6, Q-7, Q-15, Q-?, and Q-12). Objective tests were also included, but these are not of interest for the present discussion. The analysis was based on tetrachoric *r*'s and used the multiple group centroid method. Thirty-one visual rotations were made before a satisfactory simple structure was reached. Out of the 12 factors extracted, the authors claim finding a match for two BR and Q measures.

Q-13	Obsessionally careful and considerate (1)	.76
BR-G	Positive character integration (emotionally stable, objective, dependable; persevering, determined, strong-willed, thorough; responsible, trustworthy)	.66
BR-K	Socialized, cultured mind (intellectual, cultured, introspective, analytical; polished, polite, refined in dealing with people; sensitively imaginative, intuitive)	.50
BR-N	Sophistication (polished, polite, refined in dealing with people; cool, aloof, ignores people; aesthetically fastidious)	.39
Q-17	Lack of annoyance at superiority (1)	−.36

The first factor for which a match is claimed is Factor 8 which had the following loadings of BR and Q factor scores:

The authors claim a match between BR-G and Q-13. When one examines the original correlations (Saunders, 1949), it is found that BR-G and Q-13 correlate .02. The highest correlation of Q-13 with any BR measure is .22 with BR-K.

The second claim for a match is based on Factor 3 which has the following BR and Q loadings:

BR-L	Paranoid schizothymia (jealous, hostile; worrying, anxious; shy, bashful, sensitive; suspicious, broods, feels persecuted)	.41
Q-17	Common annoyance at superiority (1)	.42

In setting out some of the results below, brief descriptions of the variables entering into the behavior rating factor estimates are given in parentheses. Since each rating variable is actually a complex description, commas are used to separate adjectives within a rating variable and semicolons are used to separate rating variables. The description of only one pole of the bipolar variables is given. A fuller description of the variables can be found in Cattell (1947). The number of questions entering into each questionnaire factor estimate is given in parentheses following the factor title. The actual questions used can be found in Saunders (1949) or Cattell (1950a).

The authors have this to say about the match: "This factor exhibits not only the most precise hyperplane of any factor, but also one of the most clear-cut matches of primary factors

from each of the realms of data" (Cattell & Saunders, 1950, p. 251). The correlation between Q-17 and BR-L (Saunders, 1949) is .01. The highest correlation of Q-17 with any BR factor is −.24 with BR-K.

Three other relationships between BR and Q factors are assumed to exist, although the reported data leave much doubt as to the validity of these assumptions. Factor 4, which Cattell and Saunders assume is a second-order fusion of factors H and F, had the following loadings:

BR-H_3	Adventurous cyclothymia (sum of H_2 and H_1 below)[1]	.77
BR-H_2	Adventurous cyclothymia (marked interest in opposite sex; gregarious, sociable, likes parties)	.73
BR-H_1	Adventurous cyclothymia (composed, free from shyness)	.66
BR-D_1	Infantile sthenic emotionality (demanding, impatient, self-centered; attention-getting, shows off, brags)	.72
BR-D_2	Infantile sthenic emotionality (self-willed, egotistic, headstrong, predatory; unscrupulous, not conscientious, tends to be dishonest, selfish)	.66
BR-F	Surgency (energetic, alert, quick, spirited; cheerful, optimistic, enthusiastic; adventurous, bold, happy-go-lucky; talkative)	.64
BR-C	General emotionality (neurotic fatigue, irrationally irritable, jumpy, nervous symptoms; emotional, excited easily; changeable, undependable, moody)	−.63
Q-2	Lack of shyness (1)	−.52
BR-G	Lack of positive character integration (see factor 8 above)	−.40
Q-10	Rhathymia (5)	.40

Cattell and Saunders have the following to say about this factor:

> There is reason to believe that this represents a failure of analysis. Scattered evidence, i.e., alternative rotations of the present data, indicates that Q-2, lack of shyness, is the Q factor equivalent of (BR) H, and that Q-10 (Guilford's Rhathymia) is the match of (BR) F surgency. Here all four are involved in a single factor along with (BR) D. This almost certainly represents a second-order factor, due to the correlations among our factors being unduly high, as a result of biased estimates, or to a failure to extract sufficient factors to obtain separation of H and F (p. 251).

Cattell and Saunders present no correlations or factor plots which might serve to bolster their tenuous arguments. In the reported data there is no basis for deciding what Q factors go with what BR factors at the primary trait level. The evidence suggests at best that a second-order extraversion factor in behavior ratings may have a Q-data counterpart. However, the present associations are not consistent with those reported by Cattell later (1957) as components of the second-order extraversion factors in either BR or Q realms. Examination of the original correlation matrix (Saunders, 1949) gives no further support to the assumed matches. Q-10 correlates .26 with BR-H_3 and .20 with BR-F. Q-2 correlates −.44 with BR-H_3 and −.39 with BR-F. Q-10 has correlations of .20 or higher with BR-G, BR-F, BR-H, the highest r being with BR-G. Q-2 has correlations above .20 with *all* of the BR factor scores loading on Factor 4 as presented above, the highest r being with BR-H_3. The conclusion is necessary that this study does not give evidence for the unique linking of BR-H with Q-2 and BR-F with Q-10.

The final assumed relationship

[1] The loadings on H_2, H_3, and H_1 are spurious due to the part-whole relationships and the resulting correlated error.

between BR and Q factors was derived from the loadings on Factor 10 which were as follows:

BR-I	Sensitive, imaginative, emotionality (demanding, impatient, self-centered; dependent, immature; aesthetically fastidious; frivolous, undependable, thoughtless; sensitively imaginative, intuitive)	.43
BR-G	Lack of positive character integration (see Factor 8 above)	−.40
Q-5	Emotionally sensitive self-sufficiency (5)	.39
Q-3	Interested in understanding nature (1)	.38
BR-N	Sophistication (see Factor 8 above)	.34

Cattell and Saunders assume that Q-5 goes with BR-I, though again they present no evidence to support their selection of these particular BR and Q factors as being matched. The correlation between Q-5 and BR-I (Saunders, 1949) is .06. The highest correlation of Q-5 with any BR factor is .20 with BR-J, Neurasthenia.

In discussing the results, the following interesting and appropriate comments are made:

> If the above factorization is doing what we intended it to do only one conclusion is possible, i.e., except for two or three instances, the known personality factors, contrary to our hypothesis, are not outcrops of the same factors in different media. This does not mean that every factor may not have a manifestation in all three regions; it only means that the examples we have taken from each region do not coincide. . . . Obviously, it is the questionnaire factors that lie most in a space of their own, whereas the behavior rating factors, . . . , are rarely uninvolved (Cattell & Saunders, 1950, p. 256).

The authors then go on to point out methodological difficulties which might lead to changing this conclusion at a later date. Among these difficulties are low reliabilities, the assumption that means and sigmas on behavior ratings were approximately equivalent for men and women, and the "obtrusive" problem of second-order factors. In addition, Cattell and Saunders found that some of their communalities ran above unity and that they could not achieve satisfactory restoration of the original correlations from the factor loadings.

Whatever the methodological difficulties, and the present writer agrees that there were many, this study by itself surely cannot be taken as supportive evidence for the assertions concerning successful matching. However, this was a "pioneering study," full of methodological pitfalls. What then do the two more recent studies have to offer in support of the claim that there are secure linkages between corresponding BR and Q factors?

The second study by Cattell and Saunders (1955) turns out to be a reworking of the same data which were used in the first study, except that the analysis was limited to the 240 women. Again tetrachoric correlations were used. Fifteen factors were extracted of which 11 were "identified" with those from the first study. One new factor appeared and three were not interpreted. Adequate evaluation of this study is not possible since the writer was not able to obtain the original correlation matrix.[2] However, of the factor matches assumed to exist in the first study, only the co-loading of BR-H and Q-10 (Rhathymia) was replicated when just the women's data were used. Even this result is not firm since Q-12, masculinity, loads just as high on the H factor as Q-10

[2] Personal communication with Cattell suggested that having the *r* matrix would not add much to the picture.

(.41 and .42, respectively). Rather than detailing the confusing picture of loadings which appeared in this second study, the writer would like to quote Cattell's discussion of its results in his recent book (1957). Cattell writes:

> The second study improved on the first by keeping homogeneity of sex (240 women). Also, its design reacted to the absence of "opposite numbers" to certain factors in the first study by taking out slightly more factors in each medium, i.e., by searching for factors of lower variance. Its results were again complex, but consistent with the first, suggesting that complexity is not merely chaos or error (p. 324).

In his discussion of this study, Cattell gives the reader no indication that the same data were being reanalyzed, and he implies that the "new" study confirmed a previous one. Since the second study was based on two-thirds of the cases from the first study, some consistency of findings should be expected. Actually the results *were not consistent* with regard to the relations of L and Q factor scores. Since a new sample was not used in the second study, Cattell is hardly justified in saying that these results suggest that the first study was not merely chaos or error. At any rate, this study does not add to the evidence for a secure linkage between BR and Q factors.

The third study cited in support of the alleged linking of BR and Q factors (Cattell & Beloff, 1956) involved 153 eleven-year-old boys and girls who were rated on behavior factors C, E, F, G, and L. Four to 10 variables were used to estimate each factor. These variables were so summed as to give two equivalent estimates of each factor. The questionnaire factors involved the items in the Junior Personality Quiz (Cattell & Beloff, 1953), 7 to 13 items being used per factor. Again two estimates were made of each factor. As in the previous studies, objective test measures were included, but these are not important to the present discussion. After computing product-moment correlations, 13 factors were extracted using the group centroid method and rotated to simple structure by the quartimax (Neuhaus & Wrigley, 1954) criterion. Two additional graphic rotations were required to obtain satisfactory simple structure. When the 13 factors are examined, *in no case do both a Q factor and a BR factor load on the same factor.*

These three studies represent the sum of the evidence referenced by Cattell in support of the contention that BR and Q factors cross match. Even if the results of these studies were positive, they could not be considered conclusive with regard to the present 16 PF questionnaire factors, because of the unknown relationships of the present questionnaire factors to those used earlier.

There are two further unpublished studies from Cattell's laboratory which have some bearing on the issue. Meeland (1952) took as his doctoral thesis the problem of distinguishing between factors A, F, H, and L, all of which fall in the extraversion-introversion area. Objective test, behavior rating, and questionnaire data were obtained on 102 college men from six housing groups. Again the objective test data will not be discussed. The rating factor scores were obtained by having each member of each group rate every other member (N varied from 14 to 22) on three variables loading on each factor. The full description of the rating variables can be found in Cattell (1947). Only a brief description of each will be given below:

BR-A	Cyclothymia-Schizothymia		
a.	Easy going, good natured	vs.	spiteful, critical
b.	Adaptable	vs.	inflexible, rigid
c.	Ready to cooperate	vs.	languid, slow
BR-F	Surgency-Desurgency		
a.	Cheerful, happy-go-lucky	vs.	depressed, anxious
b.	Talkative	vs.	taciturn, introspective
c.	Energetic	vs.	subdued, languid
BR-H	Adventurous Cyclothymia-Withdrawn Schizothymia		
a.	Adventurous, likes to meet people	vs.	shy, timid, withdrawn
b.	Marked interest in opposite sex	vs.	uninterested in opposite sex
BR-L	Trustful Cyclothymia-Paranoia		
a.	Trustful	vs.	suspicious
b.	Free from jealous tendencies	vs.	jealous
c.	Socially composed	vs.	bashful

Forms A and B of the 16 PF (Cattell, 1950b) were used to gain questionnaire measures of factors A, F, H, and L.

Meeland normalized the distribution of all variables and used product-moment correlation coefficients in his analysis. The multiple group centroid method of factor analysis was used. Blind rotation was followed for 11 full rotations, and then inspection indicated that Factor 1 was a fusion of F and H. An attempt was then made to separate F and H by extracting another factor from the residuals and rotating further. Six further rotations of all factors were carried out. Two of Meeland's 11 factors showed co-loadings of behavior rating and questionnaire factors.

Factor 1 (actually a reference vector) had the following loadings:

BR-F	.73
BR-H	.65
BR-L	.56
Q-F	.47
Q-H	.44
BR-A	.43

Factor 3 had the following loadings:

Q-H	.60
BR-H	.55
BR-F	.50
BR-L	.47
Q-A	.36
Q-F	.34
Q-L	−.31

Meeland argues that Factor 1 "is best considered as F with considerable contamination with H," and that Factor 3 is H contaminated with F. When the present writer plotted the loadings on Factor 1 and 3 against each other, he found a clear-cut hyperplane nearly orthogonal to the vectors for F, H, and L which suggests that the best simple structure was contained in Meeland's original Factor 1 which fused F and H. It is obvious that his attempt to separate F and H by extracting another factor was an arbitrary procedure which led to poorer simple structure. When one looks at the correlations between the various BR and Q factors (Table 1) the reasons for his difficulty in separating these factors becomes apparent. BR-F and BR-H correlate .779, BR-F and BR-L correlate .790, and BR-H and BR-L correlate .714. Looking at the cross-media correlations for A, F, H, and L it is found that each of the Q factors correlates higher with at least one other factor which it is not supposed to match than it does with the factor it is supposed to match. Again, the possibility remains that a second-order BR extraversion factor might match with a second-order Q extraversion factor, but this study does not support the one-to-one

TABLE 1

CORRELATIONS AMONG BR AND Q FACTORS
(Meeland, 1952)

Variables	BR-A	BR-F	BR-H	BR-L	Q-A	Q-F	Q-H
BR-F	.437	—					
BR-H	.425	.779	—				
BR-L	.399	.790	.714	—			
Q-A	.106	.262	.276	.292	—		
Q-F	.180	.490	.523	.285	.139	—	
Q-H	.155	.597	.586	.516	.404	.485	—
Q-L	−.268	−.298	−.186	−.245	−.254	−.033	−.291

Note.—*N* is 102 college men. Q-L is scored for the paranoia direction while BR-L is scored for the trustful cyclothymia direction.

matching of L and Q factors at the primary trait level.

Meeland (1952) appropriately concluded that, "Generally, there was a failure of the ratings, questionnaire and objective test measures to emerge simultaneously in their appropriate factors" (p. 52).

The magnitude of the correlations between L and Q factors in this study, even though not supporting Cattell's assertions, are of interest because they are higher than have been found in most other studies. The writer would like to suggest that the special testing conditions, in which each subject first evaluated every other member of his house on the rating factors and then was given the questionnaire (to be completed within 48 hours), set up a situation wherein each self-rater was likely to be using a similar frame of reference, thus making the cross-comparisons of self-ratings more meaningful. Also, it should be noted that each subject had the rating scales in his possession a month before the actual ratings were made. Having the rating scales available might have led to closer self-examination relative to other members of the group and thus increased the validity of the self-ratings. Needless to say, such results while of theoretical interest, could not be generalized as validity data for the 16 PF under typical conditions of use.

The final study from Cattell's laboratory was carried out by Horowitz (1951) on 60 college women. Horowitz was concerned primarily with finding objective tests to measure factors A, H, and L. She included behavior ratings found previously to load on factors A, F, H, and L (nine ratings in all) and the 16 PF (first edition). Unfortunately, the only correlations reported are between the nine behavior ratings and Factor A of the 16 PF. Horowitz found correlations ranging from .43 to .79 between the 16 PF Factor A and rating variables of *adaptability*, *good naturedness*, *cooperativeness*, and *attentiveness to people* (all found previously to load BR-A). On the basis of the reported data, the reader does not know to what extent these same rating variables might correlate with other questionnaire factors, and so this evidence is of limited value. It should be remembered that with college men under similar rating conditions, Meeland found Q-A to correlate only .106 with BR-A.

In reviewing the literature, the writer was able to find only one other

study (at the primary trait level) where factored questionnaires and factored behavior ratings were used. In this study (Becker, Peterson, Hellmer, Shoemaker, & Quay, 1959), 57 mothers and 55 fathers were rated on 16 of the Fels Parent Behavior Rating Scales. These scales had been previously factored by Roff (1949). Guilford's 13 questionnaire factors were also given. Separate factor analyses were completed for mothers and for fathers, using centroid factors and the quartimax principle of rotation (Neuhaus & Wrigley, 1954). It was found that ratings of sociability loaded .62 on a mother-data factor which was primarily defined by loadings of .83 on Guilford's Rhathymia (*R*) and .82 on Guilford's Sociability (*S*). The loadings on *R* and *S* are undoubtedly spurious because many of the same items contribute to each factor. The correlations between the sociability ratings and *S* and *R* were .59 and .40 respectively. The only other factor to show appreciable loadings of rating and questionnaire scores was a father-data factor on which the Fels activeness rating loaded .69 and Guilford's General Activity (*G*) loaded .54. The correlation between the activeness rating and Guilford's *G* was .40.

Most of the Guilford factor scores "collapsed" into a single adjustment factor very similar to Cattell's second-order anxiety factor or to the second-order emotional stability factor reported by Eysenck (1953). This questionnaire adjustment factor, however, showed little relationship to ratings of adjustment or to problem behavior in the children of these parents. Becker et al., concluded that the "domain of personality being tapped by the Guilford inventories is quite limited." The reader should note that even the positive evidence for the matching of sociability in mothers (Cattell's H factor) and activeness in fathers (no Cattell parallel) is suspect because the ratings were based on a one-hour interview. These ratings were contaminated with self-reports of behavior, and it is therefore not too surprising to find some correlation with other self-report measures. The surprising and important part of this study was that there were not more relationships between the interview ratings and questionnaire scores.

Discussion

It is apparent that the present evidence does not support the claim for "secure linkage" of BR and Q factors. This does not necessarily imply that future research using more reliable and factor pure measures may not still prove Cattell's proposition to be correct. However, there are a number of reasons why the present writer considers this outcome exceedingly unlikely. Proof of Cattell's propositions requires two things: (*a*) that corresponding BR and Q factor scores be significantly correlated with each other, and that these correlations approach unity when corrected for attentuation, assuming that somehow rating bias and response bias factors have been partialed out; and (*b*) that when factored in the same matrix and rotated blindly to simple structure, the corresponding BR and Q factor measures load on the same factors and *no other factors*, except possibly rater bias or response bias factors.

The first criterion is necessary because of the inability to state the statistical significance of a factor loading and because of the attenuating effect of unreliability on factor loadings. The second criterion is necessary to demonstrate the correspondence of factor structure in

the two media. Cattell has paid little attention to the "no other factor" requirement of the second criterion in evaluating his data. If both the first criterion and the "no other factor" part of the second criterion are ignored, as was done in the Cattell-Saunders papers, one is faced with the possibility that allegedly corresponding BR and Q factor measures could load as high as .70 on the same factor when the correlation between the factor measures is zero. This possibility assumes unit communalities. However, if the communalities are only .80, each factor measure (BR and Q) could still load .57 on the same factor without being correlated with each other.

Obtaining factor measures which could meet the above requirements, even if the correspondence exists, would be exceedingly difficult with present-day techniques. Typically with cross-media studies, one finds factor measures within a given medium coalescing. This difficulty arises for several reasons. The technical skill necessary to construct factor measures which are free of other factors in the same media is lacking, even granting the added degrees of freedom permitted by the use of oblique factor coordinates. Part of this difficulty arises from response biases of various sorts which, presumably, could be minimized or controlled with further knowledge. However, the major difficulty arises from the fact that any single rating scale or questionnaire item is likely to be factorily complex, and that while this problem can be handled theoretically (i.e. by balancing an item loading positively on an irrelevant factor with an item loading negatively), in practice this problem has not been solved.

Another problem (related in some ways to the above) is the fact that with present day measurement techniques there is no way of specifying factor order in the personality domain independently of a specific, arbitrary set of variables. Since it is not possible to specify beforehand the sampling of variables in BR and Q domains appropriate to a given order, there is no way of knowing that a given set of BR and Q variables will all lead to first-order factors which can then be cross-matched. In view of the haphazard method of item sampling, it is possible that questionnaire factors will lie at different orders of generality from behavior rating factors. The relationship of order may also vary from factor to factor. Since there is no a priori solution possible to the variable sampling problem, it is doubtful if Cattell's proposition is testable outside of a given arbitrary sampling system.

Still another problem is the fact that the common language structure of a group of self-raters and judges could produce a pseudo matching of factor contents (salient variables). However, this meaning system bias would not necessarily lead to a *matching of factor scores*. The present writer believes that much of the hoped for cross-matching of BR and Q factors derives from the fact that common meaning systems bias the appearance of common patterns in "self" and "other" judgments. Such a result is, of course, artifactual and related to the problem of variable homogeneity. The greater the variable homogeneity (which in the Q-data media approaches the extreme of asking the same question again and again), the greater the probability that common meaning systems rather than common behavior patterns are producing apparent factor matches. Cattell, of course, is not insensitive to this problem (Cattell, 1957).

Beside the statistical and technical difficulties with obtaining a proof of the cross-matching of BR and Q factors, the writer finds it difficult to believe on psychological grounds that such matching should occur in random samples of the population "with all but a few factors." The proposition that BR and Q factors can be cross-matched requires that a person's judgments of his own behavior *correlate with* and show a *pattern similar to* that derived from an external judge (or judges). This is an unlikely occurrence for several reasons: (*a*) defenses of various sorts act to limit self-awareness, (*b*) the frame of reference (adaptation level) for making self-judgments will vary from person to person, making such scores not comparable, and (*c*) the experiences to be judged of necessity constitute different samples of behavior for self and other raters.

These points can be illustrated by the analogous problem of matching factors in parents' ratings with factors in teachers' ratings of children. The parent, like the self-rater is biased by his emotional involvement with his child and apt to be defensive about the child's weaknesses. Each parent makes his judgments from a different frame of reference, depending on the number of children his child's age he has known, memory ability, and the like. Finally, the parent bases his ratings on a different sample of behaviors from that of the teacher. In a study with Cattell (Peterson & Cattell, 1958), only part of which has been published, Peterson had teachers in the University of Chicago Nursery School rate 80 children, four to six years old, on 36 variables, and had parents (mostly mothers) rate the same children on a different sample of 44 variables which covered much the same behavioral domain. Peterson and Cattell were able to identify 13 factors in each analysis, named for Cattell's factors A, C, D, E, F, G, H, I, J, K, L, M, O. These two sets of factors were cross-matched in terms of psychological meaning. When the correlations between the teacher *factor scores* and the parent *factor scores* were computed for the matched factors, they were found to range from −23 to +25, averaging approximately zero. While this is an extreme example, it parallels most of the problems involved in matching BR and Q factors.

The truth on this cross-matching issue probably lies somewhere between the position taken by Cattell and the extreme example outlined above (cf., Guilford, 1959). There are undoubtedly some points where questionnaire factors and behavior rating factors overlap, but there are just as undoubtedly many points where they do not overlap.

Direction of Future Research

Psychologists are going to continue to use questionnaires because of their convenience. In terms of the present state of knowledge, it would seem mandatory that factors in personality inventories be viewed as dimensions of self-perception or dimensions of the self-concept rather than in terms of assumed behavioral dimensions. It is of course completely justifiable to find out just what kinds of behaviors factored questionnaires will predict, as both Guilford and Cattell have done. Further work along this line is in order.

The use of questionnaires is typically justified in terms of one of two rationales, namely, as a substitute for a standardized interview, or as a means for the economical assessment of personality traits. It is only the latter rationale which is of concern at the moment. Historically there have been two major criteria used in scor-

ing questionnaire items in personality assessment: (*a*) an internal consistency criterion, which has gradually given way to the more sophisticated simple structure factor criteria and (*b*) a behavior rating criterion. Cattell and Guilford have taken the first approach, Hathaway and Gough the latter. Each of these methods has its liabilities. The factor criterion gives scores which lack a predictive behavioral reference (though one might be interested in self-perceptions per se). The behavior rating criterion, as used by Hathaway and Gough, overcomes this difficulty by building predictive validity into the scoring system. Unfortunately, the resulting test scores from this latter approach typically have had high correlations among themselves, indicating that considerable time and effort were being wasted in measuring, say, the 18 variables in the Gough inventory (1957) which could probably be handled with four or five uncorrelated measures.

The writer would like to suggest a third approach to this problem which might overcome some of the above difficulties. It is suggested that *behavior rating factors* (as orthogonal as possible) be used as the criteria in developing questionnaire measures of personality. More specifically, the writer believes that Cattell's behavior rating factors (fusing a few of the highly interrelated factors and eliminating a few of the small variance factors) have been sufficiently replicated on various populations to be seriously considered as criteria for questionnaire development. Since the cultural experience of men and women in our society is quite different, it would probably be necessary to develop separate questionnaires (beginning at the item validity level) for men and women. This methodology is applicable to various age groups, or to groups of special clinical or theoretical interests (i.e. parents). Attention will, of course, have to be paid to the problems of response bias and validity of report. A final possibility of this approach is the development of a single questionnaire which can be scored in terms of either a self-perception-factor criterion or in terms of a behavior-factor criterion. The differences in the two kinds of scores (assuming adequate reliability) might provide interesting leads about possible areas of conflict or defensiveness.

Summary

This paper challenges assertions made in recent publications by Cattell and his co-workers to the effect that present research has shown in most cases a one-to-one matching of behavior rating and questionnaire factors. A review of the evidence fails to support these assertions. Many of the technical, statistical, and psychological issues involved in the possibility of matching behavior rating and questionnaire factors are discussed. Suggestions for future research using questionnaires and behavior ratings are given.

REFERENCES

Becker, W. C., Peterson, D. R., Hellmer, L. A., Shoemaker, D. J., & Quay, H. C. Factors in parental behavior and personality as related to problem behavior in children. *J. consult. Psychol.*, 1959, **23**, 107–118.

Cattell, R. B. Confirmation and clarification of primary personality factors. *Psychometrika*, 1947, **12**, 197–220.

Cattell, R. B. The primary personality factors in women compared with those in men. *Brit. J. Psychol.*, (Statis. Sect.) 1948, **1**, 114–130.

Cattell, R. B. The main personality factors in questionnaire self-estimate material. *J. soc. Psychol.*, 1950, **31**, 3–38. (a)

Cattell, R. B. *The sixteen personality factor*

questionnaire. Champaign, Ill.: Inst. for Pers. and Ability Testing, 1950. (b)

CATTELL, R. B. *Personality and motivation structure and measurement*. New York: World Book, 1957.

CATTELL, R. B. & BELOFF, H. Research origin and construction of the IPAT Junior Personality Quiz. *J. consult. Psychol.*, 1953, **17**, 6, 436–442.

CATTELL, R. B., & BELOFF, H. La structure factorielle de la personalité des enfants de onze ans à travers trois types d'epreuves. *Rev. Psychol. appl.*, 1956, **6**, 65–89.

CATTELL, R. B., & SAUNDERS, D. R. Interrelation and matching of personality factors from behavior rating, questionnaire and objective test data. *J. soc. Psychol.*, 1950, **31**, 243–260.

CATTELL, R. B., & SAUNDERS, D. R. Beiträge zur faktoren-analyse der persönlichkeit. *Zeitschrift fur experimentelle und angewandte Psychologie*, 1955, **2**, 325–357.

CATTELL, R. B., SAUNDERS, D. R., & STICE, G. *Handbook for the sixteen personality factor questionnaire*. Champaign, Ill.: Inst. for Pers. and Ability Testing, 1957.

EYSENCK, H. J. *The structure of human personality*. New York: Wiley, 1953.

GOUGH, H. G. *California psychological inventory manual*. Palo Alto: Consulting Psychol. Press, 1957.

GUILFORD, J. P. *Personality*. New York: McGraw-Hill, 1959.

HOROWITZ, JOAN. Objective personality tests investigating the structure of altruism in relation to source traits A, H, and L. Unpublished master's thesis, Univer. Ill., 1951.

MEELAND, T. An investigation of hypotheses for distinguishing personality factors A, F, and H. Unpublished doctoral dissertation, Univer. Ill., 1952.

NEUHAUS, J. O., & WRIGLEY, C. F. The quartimax method: An approach to orthogonal simple structure. *Brit. J. statist. Psychol.*, 1954, **7**, 81–91.

PETERSON, D. R., & CATTELL, R. B. Personality factors in nursery school children as derived from parent ratings. *J. clin. Psychol.*, 1958, **14**, 346–355.

ROFF, M. A. A factorial study of the Fels parent behavior scales. *Child Developm.* 1949, **20**, 29–45.

SAUNDERS, D. R. The relation of behavior rating, questionnaire and objective test personality factors. Unpublished master's thesis, Univer. Ill., 1949.

SCHEIER, I. H., & CATTELL, R. B. Confirmation of objective test factors and assessment of their relation to questionnaire factors: A factor analysis of 113 rating, questionnaire and objective test measurements of personality. *J. ment. Sci.*, 1958, **104**, 608–624.

(Received February 28, 1959)

PSYCHOLOGICAL BULLETIN
Vol. 57, No. 3, 1960

TEMPORAL EXPERIENCE[1]

MELVIN WALLACE
System Development Corporation
AND
ALBERT I. RABIN
Michigan State University

The interest of modern psychology in the experience of time ("time sense," time perception, time estimation, etc.) dates back to the last century. Reviews of experimental studies concerned with a number of aspects of time have been published in the *Bulletin* on several occasions. The last one, by Gilliland, Hofeld, and Eckstrand (1946), appeared some 10 years ago.

During the past decade or two the interest in the psychology of time has not only increased but has broadened as well. Research workers have not been satisfied to study temporal experience or "time perception" as an isolated phenomenon. A trend has developed in the direction of investigating the relationships between temporal experience and other personality phenomena, normal and abnormal.

In addition, considerable attention has been given to more general aspects of temporal experience. This tendency has shown concern with the molar rather than the molecular approach to temporal phenomena. The terms "time perspective" and "time orientation" have become familiar to social psychologists, personologists and clinicians. Workers have become involved in the study of "macro-events" revolving around the relationships between persons' past, present and future within their phenomenological frames of reference.

The present review will attempt to survey the more recent studies in time perception and methodology along traditional lines as well as studies in connection with the "new look" which involves the personality correlates of time perception. Research embracing the seemingly disparate topics of time perspective and time orientation will also be discussed. Because of the relatively uncrystallized state of this subject matter, not only experimental studies, but also the more empirical-observational, exploratory and frankly speculative papers in this area will be included.

DEVELOPMENTAL ASPECTS OF THE CONCEPT OF TIME AND TIME PERCEPTION

Though philosophers resorted to a genetic explanation of the "time sense," there is ample empirical evidence which indicates the developmental nature of this capacity. Despite Kant's speculations, there is evidence to indicate that the capacity to experience time and to estimate it is a gradually developing human characteristic (Fraisse & Vautery, 1952). Within the last few years, efforts to present formal accounts of the development of temporal concepts and experience have emanated from two major lines of approach: (*a*) psychoanalytically oriented the-

[1] The writers wish to express their appreciation to Howard H. Fink for his contribution to an earlier draft of this paper.

orizing and (*b*) extensive empirical investigation.

Psychoanalytic Views

Although not directly concerned with the problem of the development of the time sense, Freud (1950, 1952; Schneider, 1948) suggested that the notion of time was derived from the manner in which the Perceptual-Conscious system functioned. This system was considered by Freud to be the medium through which impulses passed to cathected external objects and stimulation from the outside world returned to what Freud termed "unconscious memory systems." When these cathexes were withdrawn, however, breaks in the continuity of this dynamic process occurred, leading oftentimes to a complete absence of the external world. For Freud, such periodic interruptions lay "at the bottom of the origin of the concept of time" (Freud, 1952). Implied in this formulation is the view that the development of the notion of time is a gradual process, resulting not from a single withdrawal of a cathected impulse but from many such withdrawals extending over a period of time.

At least an implicit acceptance of this view is to be found in other psychoanalytic contributions to this area. For the most part, however, these reports are primarily concerned with specifying the phases of psychosexual development during which the notion of time emerges. Yates (1938), for example, attributes the development of the appreciation of time to the patterns of body rhythm that become manifest at the oral period. These rhythmic patterns, Yates continues, become related to the periodic frustrations and satisfactions of bodily needs (especially the intake of food), thereby providing the basis for the development of the time sense. Fenichel (1945) and others (Dooley, 1941; Arieti, 1947), on the other hand, focus upon the anal phase of development as primary, and note the influence of programs of toilet training upon the acquisition of the time sense. According to other psychoanalysts, however, later phases of development are considered to be more significant in this process (Bonaparte, 1940; Oberndorf, 1941).

Theoretical formulations from such psychoanalytic sources, so far without experimental validation, allow the following reconstruction of the development of the experience of time: the young organism begins to experience some frustrations quite early. Ordinarily, however, most early experiences of infancy are nonfrustrating at first. The infant is cared for and protected from his environment. Every wish is fulfilled immediately by adults in the environment. As the organism grows and develops, minor frustrations are experienced; the usual immediate gratifications and satisfaction of needs are no longer the rule. The individual must gradually learn to cope with the environmental frustrations which result from situations which frequently require postponement of gratification. He also must learn to manipulate the environment and the elements within it for his own purposes. These processes are the stuff from which the ego is made.

Several steps in the process of postponement of gratification may be posited:

1. Immediate frustration when need is not gratified.
2. Suppression of need and of emotional reaction to the experience of frustration.
3. Development of anticipation of *later* gratification on the basis of:
   a. past experiences,
   b. "hallucination" (imagination) of gratification.

4. Projection into the future—expectancy.
5. Expectancy is possible providing the distinction between reality and nonreality experience can be made.
6. Development of symbolization—via words.
7. Development of the time sense expressed by conventional language—now, later, today, tomorrow, etc.
8. The internalization of conventional time divisions and the acquisition of the abstract notion of time as a continuous, evenly flowing substance—from the past into the future.

Empirical and Experimental Studies

To a large extent, the findings of several empirically oriented investigators are compatible with the theoretical position outlined in the preceding paragraphs. Piaget (1955), for example, suggests that the earliest experience of time stems from a sensory-motor recognition of the waiting period between feedings, a view which is similar to one alluded to previously. The child's egocentricity and his relatively poor resources for differentiating clearly between the external world and himself, Piaget maintains, interfere with the development of objective concepts of time. As this early phase is outgrown, such concepts begin to emerge although they are not verbalized until the age of four. During this silent period, however, the ability to extend the idea of time into both the past and future continues to develop. When egocentricity starts to diminish during the seventh year, the child begins to reflect upon his notions of temporality, and to compare them with the temporal notions of others. With added experience, modifications occur, and concepts are acquired which are more socialized, and thus, more adequate. A somewhat similar account of the order of development of time concepts in successive stages was reported by Bradley (1947).

Other investigators, however, tend to disagree with Piaget's formulation of the age of maturation of the notion of time. Oakden and Stuart (1922), for example, suggest that the child has little ability to conceive of the continuity and development of time until the age of 11—a view supported by Wesley (1942). According to these writers, the adult concept of time does not fully mature until the thirteenth or fourteenth year. Pistor (1939, 1940) and Bromberg (1938) also point out that the time concept develops late in childhood, but not before the age of 10 or 12. Children first understand time as being numbers on a clock, daily occurrences, etc., until finally the adult notion of time is attained.

A lack of agreement is also to be noted in many studies of the evolution of specific types of temporal concepts. In a rather detailed and systematic investigation of the emergence of the concepts of past, present, and future, Ames (1946) found that such temporal notions "come into use in a relatively uniform sequence from child to child and at about the same relative time in the life of every child." An analysis of data from direct observation and questionnaires revealed that between the age of 18 to 24 months the child lives predominantly in the present, even though some ability to project into the future may be developed during this period. "Words indicating the present come in first, then words indicating the future and finally those indicating the past. Thus 'today' (24 months) precedes 'tomorrow' (30 months) which in turn precedes 'yesterday' (36 months)" (Ames, 1946, p. 122). From two to three years of age the child begins to utilize the concepts of past, present, and future in his verbalizations, but with a greater emphasis upon the future than on the past. An increase in the number of references to these time divisions oc-

curs between age three and age four, but this period is also accompanied by some confusion between future and past activities. Nonetheless, Ames reports, this age brings with it a much greater projection into the future. At age five, the days of the week are used appropriately. At six, the four seasons are understood. The concept of the month is known at age seven, and by age eight, even extremes of the time span can be handled adequately. In general, the results of Schechter, Symonds, and Bernstein (1955) tend to support these findings.

The use of words denoting temporal concepts shows a parallel development to that of language in general. It is not until the fourth year that the concept of duration becomes apparent in the child's vocabulary. At age seven, after further acquisition of broader temporal concepts such as years and their subdivisions, seasons, etc., the notion of objective and continuous time, i.e., clock time, is acquired.

Similar observations, based on less systematic data, were made by Bromberg (1938), who states that "The development of the time sense (duration?) does not occur until after the age of 5 or 6, and develops slowly until about the age of 10 or 12" (p. 147). Another interesting observation, by the same author, which appears somewhat contradictory to that implied by Ames concerning individual differences, was: "Although the age of maturation depends on the intelligence, the formation of the time concept seems to occur at the same relative time in the life of every child. In other words, the development of abstract concepts such as time seems to meet a certain social need." Yet, it is doubtful whether the uniformity of the age at which temporal concepts appear is a defensible notion. Buck's Time Appreciation Test (1946) and its comparatively high correlation with intellectual level in addition to other findings relating time concepts to systematic instruction (Friedman, 1944, 1945) would seem to contradict this notion.

Springer's (1952) findings with older children (four to six) generally corroborate those of Ames with respect to the sequence of the development of the understanding of time. She summarizes her findings: "First, the child is able to tell time of activities which occur regularly in his daily schedule. . . . Second, the child is able to tell time by a clock . . . . Third, he is able to set the clock . . . . Fourth, he is able to explain why the clock has two hands and how each operates . . . ." (p. 95). Further support for this position with children of high I.Q., between the ages of six and seven, is presented by Farrell (1953).

It is interesting to note that there is considerable similarity between the early temporal concepts of young children and the use of time among the primitives described by Werner (1948). The young child, at first, also deals with discrete events to which he attaches certain temporal names mediated to him by culture. However, later, he acquires the abstract concept of time as duration, as continuous and evenly flowing, so characteristic of Western temporal notions, but absent in many non-Western cultures (Werner, 1948).

Conclusions differing somewhat from those found by Ames are reported in a study by Friedman (1943–1944). In testing 697 children in elementary school (including kindergarten pupils), this experimenter utilized a technique which was composed of 12 events. Each event had to be placed in one of four categories: a long time ago, a short time ago, a short time to come, and a long time

to come. It was found that the older pupils (Grades 4–6) showed a great improvement over the younger pupils in the ability to understand these indefinite concepts. On the basis of his findings, Friedman (1943–1944) concluded: "There is not so much logic in the child's thinking concerning the future as in his thinking concerning the past. . . . The idea of tomorrow appears to be less definite than is the concept of yesterday" (p. 340).

According to Eson (1951), the child of eight is primarily concerned with the present and projects into the future only on those few occasions when he is stimulated to anticipate a future event. His conception of the past is also fairly restricted, being limited to those experiences which he has had, but whose consequences are still pending. Eson also investigated what he termed "temporal emphasis" or the stress given to the past and the future in the thoughts and conversations of his *S*s. Seven groups of adults and children, representing five different age ranges, were utilized in the test of the hypothesis that the ranges of time perspective increase in the direction of both past and future in fairly equal proportions with advance in age. The *S*s were asked to list those items which they had thought about or spoken about during the two-week period preceding the interview. An analysis of some of these data in terms of its past or future reference reveals that each of the groups placed greater emphasis on the future than on the past.

Gilliland and Humphreys (1943) have compared fifth graders with adults, with respect to ability to estimate short intervals of time (9–180 seconds). The results indicate the superiority of the adults over the children by 15 to 18%. Hence, the implication here is that with respect to estimation of time, learning still takes place between preadolescence and adulthood. It seems that the internalization of the notion of time and duration in relation to objective standards is as yet incomplete in late childhood.

A more recent study (Smythe & Goldstone, 1957) found that six- and seven-year-olds are extremely variable in their estimates of one second when compared with older children, 8 to 14 years of age. Moreover, the younger children do not seem to improve nor do they learn "from specific time information," whereas the older children tend to show improvement with the aid of such information. The accuracy of estimation of one second by 14-year-olds did not differ appreciably from that of adults.

Summary

On the basis of the findings reported above, it appears that an individual's concept of time emerges early in childhood and develops gradually. By the time a child is two or three years old, he has acquired a notion, more or less limited, of a past, a present, and a future time, but until the eighth year, the child is primarily concerned with his immediate present. The time concept, with ever widening past and future references, continues to develop through the thirteenth or fourteenth year when the adult concept first emerges. At that time the notion of continuity of time and its relatively accurate estimation are reached.

The Passage of Time

A number of experiments on time perception and time estimation have been reported since the reviews by Gilliland et al. (1946) and by Woodrow (1951). As a result of considerable variation both in definition of concepts and in methodology, the results obtained are contradictory

and confusing. An analysis of methodological procedures proposed by Clausen (1950), reenforced with additional detail by Bindra and Waksberg (1956), attempts to put some order in the field of experimentation with the passage of time or duration.

Methodology

These investigators have pointed to four major methods by means of which short time intervals are judged in estimation of time experiments. These methods are: (*a*) *verbal estimation*—when *E* presents an interval (standard) and *S* is asked to respond verbally in terms of temporal units (seconds, minutes, etc.) indicating the duration of the interval; (*b*) *production*—*E* instructs *S* to produce an interval of a given length indicated verbally by *E*; (*c*) *reproduction*—*E*, by means of some operation presents an interval (standard) and *S* is expected to reproduce it as exactly as he can; and (*d*) *comparison*—*E* presents two intervals and *S* is to judge which is the longer.

Two broad methods may be distinguished with respect to the presentation of the "standard" to be judged:

1. "*Empty*" *Time.* This refers to the time or interval which is bounded by two stimuli, such as two clicks (Ross & Katchmar, 1951), two flashes of light (Abbe, 1936; Fox, 1952), and others (Bakan & Kleba, 1957). Criticism of this method stems from Kowalski (1943) who points out that "subjects tend to incorporate into the time interval these limiting sounds, which themselves take an appreciable length of time." This criticism is particularly pertinent when very brief intervals are under consideration. The study by Rodgers and Hammersley (1954) is also relevant in this connection.

2. "*Filled*" *Time.* This method is employed when the standard or interval to be estimated is started and terminated (filled) by some continuous sound or tone (Eson & Kafka, 1952; Harris, 1952; Phillip, 1944), a metronome (Cooper, 1948), light (Kowalski, 1943; Phillip, 1945), and by a variety of assigned tasks and/or activities (Coheen, 1950; Filer & Meals, 1949; Hindle, 1951; Kawasima, 1937; and others).

Clausen (1950) attempted to compare the several methods of interval presentation under conditions of filled and unfilled time. Using intervals of 5, 10, and 15 seconds, Clausen found no significant differences between the estimation of filled and unfilled intervals. He also found that the method of verbal estimation results in less accuracy (overestimation) than the other methods employed and that the method of reproduction is less consistent and less reliable. He makes the important point that the methods of verbal estimation and production deal with the relation of subjective time to "world time" (clock time) whereas the method of reproduction does not. Different functions are apparently involved.

Time Estimation Under Different Conditions

Despite the great difficulties in summarizing studies of time estimation pointed out by Bindra and Waksberg (1956), an attempt will be made to review such recent investigations in which a variety of methods of experimentation have been employed. The design of these experiments is usually simple enough, but the methods and terminology are far from being uniform.

Verbal estimation. Most of the investigations on the perception of time have utilized this method of presentation. Postman (1944) had 40 undergraduates estimate three-, five-, and seven-minute periods while engaged in several different activities (addition, cancellation, and completion). In general, there was a tendency to overestimate the actual intervals. The kind of task with which the

intervals were filled made no difference in the estimation, but a positional effect was noted. In Loehlin's (1956) study one- and four-second intervals and two- and twenty-minute intervals were employed with a similar population. The shortest and longest intervals were clearly overestimated, the four-second interval was not, whereas, the two-minute interval was overestimated. The trends were determined in terms of mean ratios of estimated to actual time. This method is different from the one in the previous study in which the accuracy of estimates was reported in terms of deviation and may partly account for the somewhat different results. It is interesting to note that Loehlin was unable to obtain a general factor of time estimation by means of his factor analysis of the data. Lhamon and Goldstone (1956) compared normals and schizophrenics on the estimation of one second, presented as a continuous tone. Both groups tended to be stable in their estimates; both groups overestimated, though the schizophrenics were more extreme in this respect. The same method was employed in another study (Smythe & Goldstone, 1957) where it was found that the tendency to "overestimate the value of 1.0 second is characteristic of all age groups." A somewhat surprising finding is reported in still another study (DE Rezende, 1950) in which intervals of 2 and 12 seconds were used. A tendency to overestimate the longer intervals and underestimate the short ones is reported and seems to be inconsistent with many of the studies previously reviewed. In a series of experiments (Roelofs & Zeeman, 1951) intervals between 400 and 3200 milliseconds were used. Order, and filled vs. unfilled conditions were considered in detail. At 3200 ms. order was more important than the contents of the interval, at 400 ms. it was less important. Empty intervals of 1800 ms. were underestimated when compared with filled ones (continuous light). No important differences were obtained between the 3200 ms. intervals and the others. Some of the practical aspects of the errors in the estimation of short intervals were considered by Culbert (1954).

Production. Eson and Kafka (1952) who noted some "diagnostic implications" in the study of time perception had their *Ss* (college students) produce 15-second and two-minute intervals under four conditions (silencing a tone, leaving a room in darkness, keeping the room lighted, and listening to a tone after switching it on). Most *Ss* underestimated the time intervals; the effect of serial position was to reduce the overestimation with succeeding judgments. The authors point out the individual differences between their *Ss* with respect to stability of estimation and the experience of the rate of passage of time. Using a 15-second interval, Falk and Bindra (1954) related greater overestimation to the experience of anxiety. A combination of the estimation and production methods is represented in Dobson's (1954) study. Intervals of 17, 38, 72 and 120 seconds were employed under filled and unfilled conditions. The effects of filled vs. unfilled conditions upon estimation are "still open to question"; also, as far as the estimation of the intervals is concerned "the tendency was for less accuracy and greater variation with increasing time intervals." An additional variation of this study is the introduction of "non-set" conditions for time estimation. Subsequent studies (e.g., Rabin, 1957)

also employed this method. Bakan (1955) found no significant differences in the estimation of a one-hour work period between set and nonset conditions.

Reproduction. This method was employed in Kowalski's (1943) experiments with intervals ranging from .48 to 16.20 seconds. He also varied the delays between intervals to be reproduced. The durations of .48 and .92 seconds were consistently overestimated, whereas the remaining intervals up to 16.20 seconds were underestimated. The final conclusion, based on analysis of variance, is that the durations of the intervals rather than those of the delays are the significant factors in under- and overestimation. This conclusion appears to be in contradiction of Fox's (1952) findings reported at a later date. Such consistent differences with respect to the stimulus were not obtained by Young and Sumner (1954) who experimented with a 30-second and five-minute interval. These investigators focus on the consistency of individuals in their reproductions of the standard rather than on the characteristic (duration) of stimuli or the intervals between stimuli during presentation. A study (Phillip & Lyttle, 1945) which attempted to relate reaction time to the ability to reproduce short time intervals (.73–1.7 sec.) yielded negative results.

Comparison. This procedure has been mainly used in the investigation of the contextual effects upon the estimation of time. These studies (Phillip, 1944, 1947; Postman & Miller, 1945) demonstrate the *anchor effect* upon the estimation of very brief intervals (usually, fractions of a second to about two seconds). Additional, more recent studies (Wright, Goldstone, Boardman, & Lhamon, 1958; Weinstein, Goldstone, & Boardman, 1958) have also demonstrated the anchor effect in temporal estimation—the "short anchor pulled judgments down and the long . . . up" (Goldstone, Lhamon, & Boardman, 1957), a finding expected in temporal estimation studies but different than that found in traditional research involving anchor effects. A study by Turchioe (1948) concludes that pronounced central tendency effects for all intervals (.78, 1.01 and 1.39 seconds) were obtained.

By means of fractionation of temporal intervals, Gregg (1951) devised a "scale of subjective time" ("temp scale"), and Ross & Katchmar (1951) using a similar method derived individual "chron-scales."

Additional Experimental Factors

Numerous investigators (e.g., Fraisse, 1948) have noted that perceptual experiences in a variety of sense modalities have their effects upon "time perception" or time estimation. Some studies (Fraisse & Oleron, 1951; Oleron, 1952) have concluded that the increase of the intensity of sound during an interval causes overestimation. Dependence of time judgment, in the young child, upon space and speed of moving objects was reported by Fraisse and Vautery (1952). The relationship of "seeming duration of time since different past events" to the "chronological time since these events" has its unique spatial representation (Cohen, Hansel, & Sylvester, 1954). Cohen (1954) also discussed the "Kappa-effect" in relation to judgment of duration. Kinesthetic cues (counting aloud) have been reported to cause the overestimation of a one-second interval (Goldstone, Boardman, & Lhamon, 1958) in children and young adults

(but not in older *Ss*). The contextual aspects of time judgment (figure-ground) were reported by Iacono (1952b) who also doubts that the very brief intervals studied are actually experiences of duration (Iacono, 1952a). Finally, the gestalt aspects of the estimation of time and its dependence upon sense modality (such as vision and audition) were illustrated by Nakajima (1951). Errors of estimates of time via visual stimuli were mostly negative, whereas the opposite is true with auditory ones. These configurational and patterning features were also emphasized in some of the investigation quoted in earlier section (Roelofs & Zeeman, 1949, 1951).

Summary

Several methods of presenting intervals for estimation in time experiments have been distinguished. Some evidence points to lesser accuracy of the verbal estimation method as compared to the others. The issue of filled vs. unfilled time still remains unsettled, though the bulk of the evidence shows little significant difference in time estimation under the two conditions. Perhaps the lack of significant results is due to the fact that the distinction between filled vs. unfilled time is in the mind of the *E* rather than in the experience of the *S*.

Generally, very brief periods (second, or fraction thereof) tend to be overestimated, while longer ones are underestimated. Positional and anchor effects in brief interval experiments have been demonstrated and have probably been confounding factors in a number of experiments which have apparently produced contradictory results. The differential effects on judgment of stimuli of different modalities, used in delineating intervals of time have been pointed out again by recent investigators. This is especially true in the judgment of brief periods the durational reality of which seems to be questioned.

The studies considered in this section have either concentrated on the stimulus conditions or deal with such conditions as a part of their investigation. One may agree with Adler (1954) that despite the recognition of the subjective factors in "time perception" many of the investigators "have not been able to free themselves from the stimulus as given and 'real'." Perhaps the greater emphasis on the phenomenological viewpoint, such as has been suggested (Cohen, 1954) will be more productive in contributing to further understanding of the processes of temporal experience or the hitherto elusive "time sense."

Passage of Time and Personality Variables

An area of recent concern to investigators has been the study of temporal phenomena in relation to a number of personality variables. Although the discussion of reports in this area may overlap to some extent with other sections, especially the one on time perspective, a sufficient amount of work has been done to merit a separate rubric.

Rapaport (1950) suggested that "time experiencing appears to be itself a phenomenon dependent on affective life," a view also supported by Clauser (1954) and Fredericson (1951). Rapaport further proposes a connection between the experience of time and temperament. Such a relationship may be exemplified in the differences among people in their readiness to act upon an idea or impulse or their tendency to delay, to postpone, to "take more time."

Several studies attempting to relate the estimation of time to some personality variables are available. The unconscious conception of parents as dominant was found to be related to overvaluation (overestimation) of periods of one minute and higher (Fisher & Fisher, 1953). An MMPI study of 71 *Ss* with high T scores on individual scales (Solomon, 1950) showed that those scoring high on one or more scales (excepting those on the manic scale) tended to make high estimates of periods ranging from 30 to 190 seconds. Also, the tendency for manic individuals to experience time more quickly than do depressed individuals was brought out. One study, mentioned earlier (Loehlin, 1956) reports an attempt to throw into the hopper measures of time estimation, questionnaire data (on phenomena of time perception) and MMPI scores, and then to factor-analyze them all. The author claims to have obtained at least "scattered information relevant to the integration of time perception into the matrix of personality." At least the information obtained *is* "scattered" since it is difficult to extricate any clear relationship between temporal experience and personality variables from the factors extracted in this investigation. However, the merits of this attempt along relatively uncharted lines cannot be denied.

Motivation as a factor in time estimation has been considered to some extent. Filer and Meals (1949) presented two experimental groups with two different tasks (writing alternate letters of the alphabet and writing more than 150 words). They were told that the task takes 10 minutes. The first group was to be dismissed from class at the end of the 10 minutes, while those of the second group who had written the 150 words in the time allowed would get a box of chocolates. The control group was told that they would return to the regular classwork upon the termination of the task. All groups were interrutped at the end of 4′ 37″ and asked to estimate the time that had elapsed. Both experimental groups overestimated the time when compared with the control group. According to the authors "an attractive goal affects the psychological distance to the goal," i.e., time passes more slowly (and is therefore overestimated) when *Ss* are motivated to have time pass. Another, related study (Burton, 1943) reports no consistent or significant findings with respect to overestimation of time under conditions of monotony (presumably, no motivation for the task but motivation for time to pass quickly). A trend in the direction of overestimation under conditions of monotony was noted, however. Hindle's (1951) investigation of "Time estimates as a function of distance traveled and relative clarity of a goal" is more clearly related to the one by Filer and Meals (1949). The findings of this investigator indicate that "during the latter portion of an activity, when it leads to a clearly defined goal, estimates of time spent increase more slowly with increments in score than estimates with no defined goal" (Hindle, 1951, p. 501). One may note that the group with "no defined goal" was subject to a greater feeling of monotony and therefore gave higher estimates. On the other hand, it would appear that these results are inconsistent with those of Filer and Meals who report greater overestimation of those who were approaching an "attractive goal." Such an attractive goal was actually not part of Hindle's complex design. The "goal" in her

experiment was mainly the completion of the task ("no reward"). Her subjects, therefore, were not so motivated to have "time pass quickly." The definition of "goal," incentive and motivation certainly vary from experiment to experiment. The results and conclusions are, therefore, not comparable. The "motivational" aspects in Hindle's study are in doubt, although the relative lack of motivation (monotony?) in the "no-goal group" may be a distinct possibility.

Under conditions of stress intervals of time tend to be overestimated according to the findings of Falk & Bindra (1954). The experimental group in this study received a shock at the end of some of the trials in producing a 15 second interval, while the control group did not. The authors conclude that "greater overestimation of the interval by the experimental group . . . is interpreted in terms of an anxious set induced by expectation of shock." A related study (Henrickson, 1948) of the "judgment of speaking time" among students suffering from various degrees of "stage fright" (anxiety?) reports no significant findings with respect to overestimation of temporal intervals.

Summary

The relatively few studies cited in this section indicate that the study of relationship between temporal experience (perception-estimation) and personality variables is still in its infancy. Variability in experimental design and terminology make it virtually impossible to compare various findings or to draw definite conclusions of a general nature from them. Some credence in a relationship between the estimation of time and stress and motivation may be justified by the findings reported in these studies.

Temporal Experience in Psychopathology

Theoretical Views

A good deal of clinical material, observational data and a review of similar reports by others concerning the disturbance of time experience in mental disorders is found in Schilder's (1936) extensive discussion of the psychopathology of time. He related particularly the phenomenon of depersonalization to disturbances in the experience of time—a loss of immediate experience of time and of the sense of the word "time," the feeling of the present as if it were the faraway past and inability to distinguish between immediate and distant past. More recently Dubois (1954) discussed the sense of time in relation to motivation of human behavior and has considered temporal distortions to be basic to all types of mental illness. Fischer (1929) maintained that there was no schizophrenic disorder which was not a space-time disorder. Others (Haranyi-Hechst, 1943; Kiersen, 1951) suggest that timelessness and other temporal distortions are quite common in schizophrenia. For Minkowski (1926), extreme distortions of subjective time were the central symptoms of schizophrenia and were prominent in other psychoses as well. Somewhat similar positions are held by Lewis (1931–1932) and by Strauss (1947) who points to temporal distortions in depressive states. Other writers (Federn, 1952; Hollos & Ferenczi, 1925) note the relationship between abnormal temporal experience and psychosis. Van Der Horst (1932) attempted to explain Korsakoff's syn-

drome as the loss of temporal aspects of experience.

According to Werner (1948) the reports of patients with temporal disturbances reveal a much greater egocentricity and concreteness in their utilization of temporal concepts than is common in the normal person. Since such temporal experiences are similar to those found in primitive peoples and children, Werner postulates that psychopathologically disturbed individuals reveal a (temporal) regression in which the normal adjustment of personal time to world time no longer is carried through completely.

Psychoanalytic writers also support the notion of regression in their discussions of disturbances in temporal concepts found in many mental disorders. In many of these views specific types of temporal distortions are associated with difficulties encountered at early levels of psychosexual development. Yates (1938) points out that prolonged disharmony between the appearance of infantile needs and their satisfaction especially in the oral period, may give rise to a loss both of the awareness of the passage of time and of sense of reality. Other writers (Bergler & Roheim, 1946; Scott 1948; Schneider, 1948) also note the close connection between disturbances in the appreciation of reality and temporal distortions. Fenichel (1945) and others (Schilder, 1936; Dooley, 1941; and Arieti, 1955) suggest that fixations at the anal level, said to be primary in the later development of obsessive-compulsive neurosis and paranoid schizophrenia, also are characterized by particular disturbances in temporal conceptions. Oberndorf (1941) discusses temporal distortions originating at both the oral and anal periods of personality development. Temporal difficulties in neurosis are also noted by Wolff (1950).

Summary. In general, the preceding accounts agree in suggesting that (*a*) disturbances in temporal experience are related to emotional or neurological difficulties, and (*b*) at least some types of psychopathology are characterized by specific temporal disturbances. This view is perhaps best summarized by Schneider (1948), who states rather generally that "the manner in which a person handles time . . . is very closely linked to the structure of his character and to the nature of his neurosis." This view is also supported by Meerloo (1950).

Empirical Findings

For the most part, little empirical validation of any of the preceding views is reported in the literature.

In a well-conceived study of time perception in neurotic, schizophrenic, and "normal" *S*s, Dobson (1954) found no significant differences among the mean estimates of 17, 38, and 72 seconds, although the schizophrenics revealed a greater variation in their estimates. Johnston (1939) reported that schizophrenics differed from normals notably in their estimates of 5-, 10-, 90-, and 100-second intervals but also in the variability of their estimates; they underestimated the short intervals and overestimated, more than the normals, the longer ones. Rabin (1957) also found significant differences (bimodal distribution) between schizophrenics and nonpsychotics in judgments of longer periods of time (one-half-hour to an hour). De La Garza and Worchel (1956) reported that schizophrenics were significantly poorer than the normals on time-orientation tests. Adler (1954) reported that schizophrenics, as compared to psychopaths

and normals, tended to overestimate short periods of time (30 seconds, minute, 2½ minutes, and 3 minutes) spent on a number of tasks. Lhamon and Goldstone (1956) demonstrated greater overestimation of one second by schizophrenics than by normals. Also it has been found that schizophrenics are less subject to remote anchor effects (Weinstein, Goldstone, & Boardman, 1958; Wright, Goldstone, Boardman, & Lhamon, 1958) than are normals.

Coheen (1950) reported that the degree of agnosia in organic patients was roughly proportional to degrees of deterioration. Fraisse (1952), comparing functional psychotics, organic psychotics and aphasics, found that patients with functional psychosis appear to have the greatest difficulty in perceiving and reproducing a short interval of nonfilled time.

Brower and Brower (1947) studied the relationship between temporal judgment and social competence in the feeble-minded and reported that mental ability has more influence on time orientation and estimation than social competence. A close relationship between the knowledge of time and the Stanford Binet in defective children was obtained in another study (Gothberg, 1949). Both studies provided evidence concerning the relative primitiveness of the concept of time in feeble-minded children.

Summary

Although a number of studies of distortions in temporal experiences as revealed in mental disturbance have been reported in recent years, the total number of investigations in this area is relatively small. For the most part, little light has been shed on the nature of the processes which underlie the manifestations of temporal distortions in psychopathological conditions. Considerable future work, both in terms of theoretical formulation and experimentation will be necessary before substantial progress can be reported in this area.

Temporal Experience Under Special Conditions

Other research efforts concerned with temporal phenomena have involved the study of temporal experience under a variety of conditions. Included among the topics investigated are changes (distortions) in temporal experience as a result of the effect of various drugs and hypnosis.

Drugs

A perusal of the literature reveals that the number of research reports in this area, since the review by Gilliland et al. (1946), is rather small. This review reported the extreme tendency to underestimate and overestimate time under mescal intoxication. A similar effect of marijuana was also reported. Quinine and alcohol caused underestimation, while caffeine and thyroxin yielded inconsistent findings.

More recently, nitrous oxide was shown to influence subjective impressions of changes in rate of passage of time with *Ss* revealing both quickening and slowing experiences (Steinberg, 1955). Another study with marijuana, but with no new conclusions, was reported (Bromberg, 1941). Neither has LSD (Lysergic Acid Diethylamide), which is supposed to produce schizophrenic-like symptoms, yielded clearcut effects upon the estimation of short time intervals (Boardman, Goldstone, & Lhamon, 1957). Another mescal report by Wolfe (1952) added little to the findings previously reported.

Hypnosis

A review of studies of the perception of time under hypnosis was published by Loomis (1951) who concluded that hypnotic subjects were sensitive to temporal factors and were able to estimate time more accurately than nonhypnotized persons. Cooper (1948, 1950) reported that the time sense may be deliberately distorted under hypnosis so that subjects may report even extremely brief intervals as long ones. Cooper and co-workers also reported that temporal distortion is found in subjects performing tasks involving motor (Cooper & Tuthill, 1952) and nonmotor (Cooper & Rodgin, 1952) learning. Other reports (Cooper & Erickson, 1950; Cooper, 1952; Eisenbud, 1947; Gross, 1949) have suggested that temporal distortion is found in dreams as well.

Summary

There seems to be some evidence that a number of drugs have a distorting effect upon the experience of time. However, the effects do not appear to be clear-cut or unidirectional in all instances. Hypnosis and dream states also have their effects. Hypnotic suggestion may give direction to the distortion. As to dreams, subjective reporting is the only source of evidence and may be fallible.

Physiological Correlates of the Experience of Time

Theoretical Views

A major focus in this area has involved efforts to specify an organic or physiological locus for the "time sense." Contributions in this regard range from the view of Coheen (1950) who states that "the structure of the temporal organization . . . is not localizable to some one portion of the brain" to directly antithetical positions espoused by Schilder (1936), that the parietal lobe and adjacent parts of the temporal lobe are the specific site of the experience of time, and Campbell (1954) who suggests that the organization of the cerebral cortex can be divided into past and future orientation.

Other writers have pointed to the association between time sense and other physical and physiological factors. A number of reports in this area have been reviewed by Gilliland et al. (1946). Davidson (1941) proposes that the development of the time sense parallels the process of myelinization. Hoagland (1933, 1934, 1943) postulated the existence of a chemical brain clock in which the judgment of time is attributable to a master chemical reaction involving the respiration of certain parts of the brain, a view also held by Piéron (1936). Others propose that time perception is based upon body metabolism and physiological factors, and Lecomte Du Nuoy (1937, 1939, 1944) has suggested quantiative biological methods of measuring the individual time sense. MacKenzie and Munsterberg, as reported by Eson (1951), also suggest that time perception is based upon rhythmical physiological processes.

Empirical Findings

On the basis of evidence available until 1946, Gilliland, et al. (1946) concluded that physiological hypotheses were inadequate as an explanation of the perception of time and that the most significant factors in the estimation of time were external. The number of publications in this area subsequent to 1946 has been negligible and has presented no experimental evidence to refute Gilliland's conclusion. It seems apparent,

however, much research remains to be done before a definitive theoretical framework can be developed to account for the relationships between internal physiological factors and external factors in the estimation of time.

The Concept of Time Perspective

Theoretical Considerations

In addition to a rather sizable number of studies regarding temporal concepts such as time sense and perception of time, a recent focus of attention has involved what has been termed *time perspective*. As far as can be determined from a perusal of the literature, Frank (1939) was the first to present a detailed consideration of this idea and to speculate about its relation to human conduct. According to this view, culturally determined attitudes about notions of temporality constitute one major aspect of the influence of culture upon behavior. He pointed out that temporal factors not only figure prominently in demarcating culturally acceptable times of occurrence for specific life-phenomena but also underlie, albeit implicitly, the transmission of cultural standards and mores, since parental child-training procedures, serving to accomplish this end, are subtly embedded in a temporal dimension. In addition to including such considerations within the framework of an explanatory formulation of the development of time perspective, Frank also described, within the context of a field-theoretical approach, the dynamic interplay between acquired conceptions of the past and the future, especially in terms of their everchanging influence upon notions of the present—a topic also discussed by Perls (1947), Mowrer and Ullman (1945), and others.

In an approach which may also be characterized as field-theoretical, Lewin (1946, 1951) was concerned with time perspective as a dimension of the "life space" and refers briefly to a formulation which might account for its development. In another reference to this topic, Lewin (1942) suggests that the social environment in which an individual lives may materially influence his time perspective—a view also discussed by Smith (1952). If a person lives in a social environment which is autocratic, for example, the future is decided and structured by the leader. An individual in a "democratic" group, on the other hand, takes an active role in determining both his own and the group's plans. In a somewhat related view, Hulett (1944) proposes that group experiences in general provide contexts for learning to anticipate the social roles of others, and therefore have an important bearing upon individual time perspectives. Ketchum (1951) also examines the relation between temporal factors and social perception.

For Rapaport (1946), the notion of anticipation is associated with a sequence of events extending into future and as such is intimately related to time perspective and the general problem of time experience. Arieti (1947) also discusses anticipation and its role in personality development and psychopathology.

Research Findings. Eson (1951) investigated the emphasis given to the past and the future in five groups of subjects of different ages. Each subject was asked to determine the temporal focus of at least 25 items derived from thoughts and conversations during the previous two-week period. On the basis of the data obtained in this manner, Eson concluded that each of the five groups tested revealed

an orientation toward the future.

Other measures of time perspective have been derived from data obtained by means of questionnaire and interview techniques. In one study concerned with the relationship between time perspective and morale in a group of prison inmates, Farber (1944) selected 28 variables from a wealth of case study and interview material. By means of correlational techniques, the relationships among these factors were evaluated, and it was concluded that the degree of pain and suffering in the prison situation correlated most highly with "factors involving time perspective, particularly the future outlook . . . ," and not with the actual length of sentence or time already served—a finding related to that reported by Field (1931). On the basis of two other studies, involving college students and also utilizing questionnaire and other direct information-getting techniques, Farber (1951, 1953) reported that an individual's current mood is more influenced by the psychological future time perspective than the situation in which he finds himself. A somewhat related finding was reported by Israeli (1933–1934) on the basis of questionnaire material obtained from college students. Israeli (1934, 1935a, 1935b) also investigated the time perspective of unemployed miners, mental patients and others, and found by means of interview rating forms that the future time perspective of the unemployed was similar to that found in the mentally ill. In passing it is to be noted that the latter investigator has introduced the method of future autobiography (Israeli, 1936) and other techniques (Israeli: 1932, 1933–34) which may also be utilized as a direct means of examining those components of time perspective which involve the future.

In attempts to obtain data regarding time perspective which might be less subject to conscious control, several techniques, characterized by relatively unstructured or projective stimulus situations, have been utilized. LeShan (1952), for example, employing a Tell-Me-A-Story technique, found evidence to support the hypothesis that the time perspectives of middle-class children are longer than those which characterize lower-class children. Other investigators (Barndt & Johnson, 1955) report that delinquent boys reveal shorter spans of time perspective and are more present-oriented than nondelinquents on a story-completion technique. A later investigation (Davids & Parenti, 1958) with the same story completion method found no differences between disturbed and normal 11-year-olds with respect to time perspective ("present oriented"). The normal and disturbed children in this study were more oriented in the present than the normal adolescents in the Barndt & Johnson (1955) investigation but were not different from the adolescent delinquent group in this respect. Expansion of the time perspective may well be one aspect of maturation.

Using a modification of the Thematic Apperception Test (TAT) as one measure of temporal orientation, Fink (1953) demonstrated that aged institutionalized individuals revealed a time perspective which was more concerned with the past than that found in their noninstitutionalized contemporaries. Bonier and Rokeach (1957) computed the percentages of past, present, and future times used in stories given by *Ss* to five TAT cards in a study of the relationship between dogmatism and time perspective and reported that *Ss* high in dogmatism, compared with low scorers on this variable, gave fewer present

and significantly more future responses. In a study of future time perspective in schizophrenia, Wallace (1956) defined time perspective as "the timing and ordering of personalized events." Two aspects of future time perspective, termed *extension* and *coherence*, were measured by means of a task in which *S* spontaneously presented a number of anticipated events and the possible age of occurrence of each of these events. *Extension* was defined as the length of the future time span which is conceptualized, while *coherence* refers to the degree of organization of the events in the future time span. A similar kind of conceptualization is provided by Arieti (1955), who discusses the "restriction of the psychotemporal field" and "seriation functions" in schizophrenic disturbances. Cohen et al. (1954) were involved in a study of what might be termed extension and coherence for past events. Teahan (1958) reported that children high in academic achievement revealed both greater reference to the future and more extensive future time perspectives than low achievers. Optimism was also related positively to future extension in this study.

In another recent paper Levine and Spivack (1957) reported on the relationship between time perspective and the ability to achieve scholastically despite the removal of an immediate incentive. These investigators employed many of the items of Wallace's (1956) method in measuring what they termed "time conception."

Future Time Perspective and Psychopathology

Theoretical views. A variety of writers suggest that emotional disturbances may be characterized by inability to conceptualize the future. Greenacre (1945) characterizes the psychopath as lacking time perspective. Strauss (1947) states that depressed patients typically are unable to abandon the past and advance toward the future. According to Eissler (1952), depressed patients do not experience the present as a bridge between the past and future, but ruminate and brood and view the future as if it were a "shapeless gap." Lewis (1931–1932) reports that cases characterized by feelings of depersonalization and a sense of unreality reveal an inability "to look into the future or to anticipate a future for oneself." Meerloo (1948, 1950), in focussing upon neurotic difficulties which involve conceptions of the future, calls attention to "anticipation neurosis." Such conditions, Meerlo suggests, are related not so much to the specific nature of childhood conflicts and early traumatic experiences, as to repeated expectations that similar unpleasant experiences will occur again.

Empirical findings. Of the many published accounts concerning temporal disturbances in mental disorder, only a few appear to have some relevance to the concept of time perspective. In one of these accounts, which is concerned with the level of aspiration of one psychotically depressed case, Escalona (1940) mentions that the patient's time perspective had become more limited and that his outlook was determined almost completely by his psychological present. Utilizing questionnaire and interview material to study future outlook, Israeli (1936) found drive with respect to future goals and the nature of anticipated possibilities served as a basis for differentiating normals and abnormals, melancholics and paranoid schizophrenics. Psychopathological groups were characterized by limited time perspectives "scarcely extending beyond the present or an attenuated perspective

comprising barely a few future possibilities" (p. 118). Wallace (1956) reported that future time perspective is influenced by the schizophrenic process to such an extent that both the length (extension) of the future span of time and the organization of its contents (coherence) are significantly reduced in patients as compared to normals. Adler (1954), in addition to the estimation of time experiments discussed in a previous section of the present review, devised an 81-item time perspective questionnaire. Many of the items differentiated to a statistically significant degree between the normals, psychopaths and schizophrenics who served as *Ss* in his investigation.

Findings on the relationship between future time perspective and anxiety, which appear contradictory to results cited earlier, were reported by Lipman (1957). Although Lipman did not study psychiatric patients but college students with high scores on the Manifest Anxiety Scale (based on the MMPI), he confirmed the hypothesis that anxiety " . . . entails experimental components that are strongly future-oriented in nature." A factor analytic treatment of the nine scales related to future time perspective yields two independent dimensions: "dismal unclarity" and "exaggerated goal frustration fears."

Summary and Critique

The investigations of time perspective, reviewed above, reveal a trend toward significant findings. Before any generalizations based upon these results can be made, however, several considerations must be taken into account. In the first place, each of the studies of time perspective cited earlier contain a somewhat different interpretation of this concept. In the reports of LeShan (1952), Bonier and Rokeach (1957), and Barndt and Johnson (1955) the concepts of time perspective and time orientation are used interchangeably, with the meaning of neither concept being explicitly specified. Both concepts are also mentioned in investigations by Eson (1951) and Fink (1953) but are utilized in somewhat discrete fashion, even though no specific attempt to indicate their possible interrelationships is made. Teahan (1958) makes a distinction between time orientation and time perspective in his hypotheses concerning the relation of these variables and academic achievement. It is to be noted that the work of Eson (1951) and Fink (1953) in this area stems to a large extent from the theoretical position of Lewin who defined time perspective as "the totality of an individual's views of his psychological past and psychological future existing at a given time" (1951). Such a formulation illustrates some of the difficulties encountered in conceptual and methodological approaches to the study of time perspective, in that the terms included within the definition require more precise specification. Concepts such as psychological past and psychological future, for example, are subject to many interpretations, and the results of any empirical investigation of time perspective derived from Lewin's formulation are ambiguous unless adequate definitions of these terms are presented. It is to be noted, moreover, that the precise meaning of any empirical findings would be dependent upon the particular set of definitions that were utilized.

Another obstacle in the way of adequate research in this area is the difficulty in selecting an appropriate methodology. To some extent the

specific definitions chosen for any interrelated set of concepts describing time perspective would limit the methodological tools that could be employed. Even within the range of potentially utilizable techniques, however, the choice and utilization of any device would in itself tend to restrict the meaning of the concept under study. As Pratt (1948) has pointed out, any fact is a function of the method by which it is derived.

In view of the above considerations, the studies discussed in the foregoing review of the literature regarding time perspective might most logically be evaluated in terms of the methods, as well as the specific set of definitions, that were employed. Both the particular methodological techniques and the specific definitions of time perspective utilized in these investigations were highly diversified. To a large extent, therefore, the results reported are not comparable, and attempts to draw broad generalizations might be seriously questioned. Nonetheless, in view of the relatively sizeable number of significant findings reported, it does seem appropriate to conclude that both the direct and indirect approaches to the problem of time perspective have been quite fruitful. Further systematic research developed from a consistent theoretical and methodological point of view appears necessary to provide a basis for obtaining additional definitive information about the nature of time perspective and its relation to phenomena in other areas.

General Summary—Time Perception and Time Perspective

The concept of time is a complex one. It ranges from relatively primitive notions of "before" and "after," as in early infancy, and definite points or events, as in preliterate cultures, to the abstract idea of a continuous flow which is divisible into equal units of duration. The latter idea, of course, is related to the observed periodicity in cosmic changes and movement.

Attempts to localize the "time sense" cortically or in any functional physiological system of the body have so far received little experimental support. There appears to be more evidence which points toward the development of temporal experience which parallels the evolvement of self or ego. This development begins early in life and continues with the increasing consciousness and discrimination of events which serve as the boundaries for the conceptualized periods of time. The relationship of the personal experiences and events to conventional units of time is learned and depends upon the cultural setting from which a person originates. The abstraction of time as continuously flowing which is boundless both in the past and future is a relatively late acquisition of the developing individual even in the most literate societies.

The capacity to estimate, reproduce or produce specific units of time is a judgmental process (Woodrow, 1951) which is based on the experience and consciousness of events and a reflection upon them in relation to previous experiences as related to conventional units of time. Like most judgmental processes, the judgment of time becomes impaired under a variety of conditions involving drugs, cortical changes and psychosis. Other, special conditions, such as extreme affective states and anxiety, may have similar impairing effects. The possibility should be considered that certain individual differences in

the awareness and judgment of time exist and are due to certain developmental and experiential events in the organism. This relationship, however, between temporal experience and personality factors is yet to be clearly demonstrated.

A further expansion of the abstraction of time is reflected in the concepts of time orientation and, particularly, "time perspective." Whereas the experiments in the perception and estimation of time deal with relatively brief periods of time (usually seconds and minutes), time perspective is concerned with long periods of time, the limit of which is one's expected lifespan itself. The units of time, in addition to direction (e.g. past, future, etc.), are usually days, weeks, months, years, and decades. Here the time periods are not artificially limited by *E*'s production of stimuli; the "points" in time which delimit the periods are important events in the life of each *S*. It follows, therefore, that time perspective involves a molar rather than molecular (or atomistic) approach to the problem of temporal behavior. It involves the total personality, memory for past events, and hopes, aspirations, and anticipations of future events. The data which are obtained in a "projective" fashion involve temporal quantity ("extension") as well as orderly arrangement of events in a logical sequence (coherence), and direction or temporal "orientation" (whether primarily past, present, or future). Finally, traditionally in perception experiments dealing with time, the stimulus or the conditions under which the stimuli were presented were the major focus of *E*'s attention. In time perspective, intra-individual, rather then extra-individual conditions, are of main concern. The person's "projection of the self in the temporal dimension," as part of the uniqueness of personality, is of major interest in the study of time perspective.

It will be, of course, of particular concern to future investigators to work out in detail a theory of temporal behavior and subsequently relate it to a variety of significant personality variables. The more precise relationship between time perception and time perspective, although delineated conceptually, is in need of empirical and experimental underpinning and substantiation. This is a task for the future.

REFERENCES

ABBE, M. The spatial effect upon the perception of time. *Jap. J. exper. Psychol.*, 1936, **3**, 1–52.

ADLER, N. The perception of time as a function of self-organization. Unpublished doctoral dissertation, Univer. of Calif., Berkeley, 1954.

AMES, L. The development of the sense of time in the young child. *J. genet. Psychol.*, 1946, **68**, 97–125.

ARIETI, S. The processes of expectation and anticipation. *J. nerv. ment. Dis.*, 1947, **106**, 471–481.

ARIETI, S. *Interpretation of schizophrenia.* New York: Bruner, 1955.

BAKAN, P. Effect of set and work speed on time estimation. *Percept. mot. Skills*, 1955, **5**, 147–148.

BAKAN, P., & KLEBA, F. Reliability of time estimates. *Percept. mot. Skills*, 1957, **7**, 23–24.

BARNDT, R. J., & JOHNSON, D. M. Time orientation in delinquents. *J. abnorm. soc. Psychol.*, 1955, **51**, 343–345.

BERGLER, E., & ROHEIM, G. Psychology of time perception. *Psychoanal. Quart.*, 1946, **15**, 190–207.

BINDRA, D., & WAKSBERG, HELENE. Methods and terminology in studies of time estimation. *Psychol. Bull.*, 1956, **53**, 155–159.

BOARDMAN, W. K., GOLDSTONE, S., & LHAMON W. T. Effects of lysergic acid diethylamide

(LSD) on the time sense of normals. *Arch. neurol. Psychiat.*, 1957, **78**, 321–324.

BONAPARTE, M. Time and the unconscious. *Int. J. Psycho-Anal.*, 1940, **21**, 427–468.

BONIER, R., & ROKEACH, M. Open and closed belief systems in relation to time perspective. *Amer. Psychologist*, 1957, **12**, 377. (Abstract)

BRADLEY, N. The growth of the knowledge of time in children of school age. *Brit. J. Psychol.*, 1947, **38**, 67–78.

BROMBERG, W. The meaning of time for children. *Amer. J. Orthopsychiat.*, 1938, **8**, 142–147.

BROMBERG, W. Marihuana intoxication. *Amer. J. Psychiat.*, 1941, **91**, 303–305.

BROWER, JUDITH F., & BROWER, D. Relation between temporal judgment and social competence in the feebleminded. *Amer. J. ment. Defic.*, 1947, **51**, 619–623.

BUCK, J. N. The time appreciation test. *J. appl. Psychol.*, 1946, **30**, 388–398.

BURTON, A. A further study of the relation of time estimation to monotony. *J. appl. Psychol.*, 1943, **27**, 350–359.

CAMPBELL, J. Functional organization of the central nervous system with respect to orientation in time. *Neurology*, 1954, **4**, 295–300.

CLAUSEN, J. An evaluation of experimental methods of time judgment. *J. exp. Psychol.*, 1950, **40**, 756–761.

CLAUSER, G. *Die Kopfur: das automatische erwachen.* Stuttgart, Germ.: Ferdin & Enke, 1954.

COHEEN, J. Disturbances in time perception in organic brain disease. *J. nerv. ment. Dis.*, 1950, **113**, 121–129.

COHEN, J. The experience of time. *Acta psychol.*, 1954, **10**, 207–219.

COHEN, J., HANSEL, C. E. M., & SYLVESTER, J. An experimental study of comparative judgments of time. *Brit. J. Psychol.*, 1954, **45**, 108–114.

COOPER, L. F. Time distortion in hypnosis. *Bull. Georgetown Univer. Med. Center*, 1948, **1**, 214–221.

COOPER, L. F. Time distortion in hypnosis with a semantic interpretation of the mechanism of certain hypnotically induced phenomena. *J. Psychol.*, 1952, **34**, 257–284.

COOPER, L. F., & ERICKSON, M. H. Time distortion in hypnosis, II. *Bull. Georgetown Univer. Med. Cent.*, 1950, **4**, 50–68.

COOPER, L. F., & RODGIN, D. W. Time distortion in hypnosis and nonmotor learning. *Science*, 1952, **115**, 500–502.

COOPER, L. F., & TUTHILL, C. F. Time distortion in hypnosis and motor learning. *J. Psychol.*, 1952, **34**, 67–76.

CULBERT, S. Systematic error in the estimation of short time intervals. *J. crim. Law Criminol.*, 1954, **44**, 685–688.

DAVIDS, A., & PARENTI, ANITA N. Time orientation and interpersonal relations of emotionally disturbed and normal children. *J. abnorm. soc. Psychol.*, 1958, **57**, 299–305.

DAVIDSON, C. M. A syndrome of time agnosia. *J. nerv. ment. Dis.*, 1941, **94**, 336–337.

DE LA GARZA, C. O., & WORCHEL, P. Time and space orientation in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, **52**, 191–194.

DE REZENDE, MARIA. An experiment on the perception of time. *Arqu. Brasil. Psicotecnica*, 1950, **2**(4), 40–55.

DOBSON, W. R. An investigation of various factors involved in time perception as manifested by different nosological groups. *J. genet. Psychol.*, 1954, **50**, 277–298.

DOOLEY, LUCILLE. The concept of time in defence of ego integrity. *Psychiatry*, 1941, **4**, 13–23.

DUBOIS, F. S. The sense of time and its relation to psychiatric illness. *Amer. J. Psychiat.*, 1954, **111**, 56–51.

EISENBUD, J. The dreams of two patients in analysis. *Psychoanal. Quart.*, 1947, **16**, 39–60.

EISSLER, K. R. Time experience and the mechanism of isolation. *Psychoanal. Rev.*, 1952, **39**, 1–22.

ESCALONA, SIBYLLE K. The effect of success and behavior upon level of aspiration and behavior in manic-depressive psychosis. *Univer. Ia. Stud. Child Welf.*, 1940, **16**, 197–302.

ESON, M. E. An analysis of time perspectives at five age levels. Unpublished doctoral dissertation, Univer. of Chicago, 1951.

ESON, M. E., & KAFKA, J. S. Diagnostic implications of a study in time perception. *J. genet. Psychol.*, 1952, **46**, 169–183.

FALK, J. L., & BINDRA, D. Judgment of time as a function of serial position and stress. *J. exp. Psychol.*, 1954, **47**, 279–282.

FARBER, M. L. Suffering and time perspective of the prisoner. *Univer. Ia. Stud. Child Welf.*, 1944, **20**, 153–227.

FARBER, M. L. The Armagedon Complex. *Pub. opin. Quart.*, 1951, **15**, 217–224.

FARBER, M. L. Time-perspective and feeling tone: A study in the perception of the days. *J. Psychol.*, 1953, **35**, 253–257.

FARRELL, MURIEL. Understanding of time relations of five, six and seven-year old children of high IQ. *J. educ. Res.*, 1953, **46**, 587–594.

FEDERN, P. *Ego psychology and the psychoses.* New York: Basic Books, 1952.

FENICHEL, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.

FIELD, H. E. The attitudes of prison inmates. *Amer. J. Orthopsychiat.*, 1931, **1**, 487–500.

FILER, R. J., & MEALS, D. W. The effect of motivating conditions on the estimation of time. *J. exp. Psychol.*, 1949, **39**, 289–336.

FINK, H. H. *The relationship of time perspective to age, institutionalization and activity.* Unpublished doctoral dissertation, Michigan State Univer., 1953.

FISCHER, F. Zeit struktus und Schizophrenie. *Z. neurol. Psychol.*, 1929, **121**, 544–574.

FISHER, S., & FISHER, RHODA L. Unconcious conception of parental figures as a factor influencing perception of time. *J. Pers.*, 1953, **4**, 496–505.

FOX, G. The effects of interval and duration of visual stimuli upon the time error. *Psychol. Newsltr.*, 1952, **45**, 1–17.

FRAISSE, P. Constant errors in the reproduction of short intervals of time. *Arch. Psychol., Geneve*, 1948, **32**, 161–176.

FRAISSE, P. Time relationships and their pathological disintegrations. *Encephale*, 1952, **41**, 122–142.

FRAISSE, P., & OLERON, G. The perception of the duration of a sound of increasing intensity. *Année. Psychol.*, 1951, **50**, 327–343.

FRAISSE, P., & VAUTERY, P. Perception of space, speed and time in the child, five years of age. II Time. *Enfance*, 1952, **5**, 102–119.

FRANK, L. K. Time perspectives. *J. Soc. Phil.*, 1939, **4**, 293–312.

FREDERICSON, E. Time and aggression. *Psychol. Rev.*, 1951, **58**, 41–51.

FREUD, S. *Beyond the pleasure principle.* New York: Liverwright, 1950.

FREUD, S. *Collected Papers*, Vol. 5, London: Hogarth, 1952, Pp. 175–180.

FRIEDMAN, K. C. Time concepts of elementary school children. *Elem. Sch. J.* 1943–1944, **44**, 337–352.

FRIEDMAN, K. C. Time concepts of junior and senior high school pupils and of adults. *Sch. Rev.*, 1944, **52**, 233–238.

FRIEDMAN, K. C. A time comprehension test. *J. educ. Res.*, 1945, **39**, 62–68.

GILLILAND, A. R., HOFELD, J. B., & ECKSTRAND, G. Studies in time perception. *Psychol. Bull.*, 1946, **43**, 162–176.

GILLILAND, A. R., & HUMPHREYS, D. W. Age, sex, method and interval as variables in time estimation. *J. genet. Psychol.*, 1943, **63**, 123–130.

GOLDSTONE, S., BOARDMAN, W. K., & LHAMON, W. T. Kinesthetic cues in the development of time concepts. *J. genet. Psychol.*, 1958, **93**, 185–190.

GOLDSTONE, S., LHAMON, W. T., & BOARDMAN, W. K. The time sense: Anchor effects and apparent direction. *J. Psychol.*, 1957, **44**, 145–153.

GOTHBERG, LAURA C. The mentally defective child's understanding of time. *Amer. J. ment. Defic.*, 1949, **53**, 441–455.

GREENACRE, PHYLLIS. Conscience in the psychopath. *Amer. J. Orthopsychiat.*, 1945, **45**, 495.

GREGG, L. W. Fractionation of temporal intervals. *J. exp. Psychol.*, 1951, **42**, 307–312.

GROSS, A. Sense of time in dreams. *Psychoanal. Quart.*, 1949, **18**, 466–470.

HARANYI-HECHST, B. Zeitbewusstein und schizophrenie. *Arch. Psychol.*, 1943, **116**, 287.

HARRIS, J. D. The decline of pitch discrimination with time. *USN Submar. Med. Res. Lab. Rep.*, 1952, No. 186, 11(3).

HENRICKSON, E. H. A study of stage fright and the judgment of speaking time. *J. appl. Psychol.*, 1948, **32**, 532–536.

HINDLE, HELEN M. Time estimates as a function of distance travelled and relative clarity of a goal. *J. Pers.*, 1951, **19**, 483–501.

HOAGLAND, H. The physiological control of judgments of duration: Evidence for a chemical clock. *J. genet. Psychol.*, 1933, **9**, 267–287.

HOAGLAND, H. The physiological control of judgments of duration. *Amer. J. Physiol.*, 1934, **109**, 54.

HOAGLAND, H. The chemistry of time. *Sci. Monthly*, 1943, **56**, 56–61.

HOLLOS, S., & FERENCZI, S. *Psychoanalysis and the psychic disorder of general paresis.* New York: Nerv. Ment. Dis. Publ., 1925.

HULETT, J. E. The person's time perspective and the social rule. *Soc. Forces*, 1944, **23**, 155–154.

IACONO, G. Contributions to the study of the perception of brief temporal intervals. *Cont. Lab. Psicol., Univer. Sacro. Cuore, Milano*, 1952, Ser. 15, 74–98. (a)

IACONO, G. An aspect of the psychological program of time perception: The direct awareness of duration. *Cont. Lab. Psicol., Univer. Sacro Cuore, Milano*, 1952, Ser. 15, 110–133. (b)

ISRAELI, N. Wishes concerning improbable future events: Reactions to future. *J. appl. Psychol.*, 1932, **16**, 584–588.

ISRAELI, N. Measurement of attitudes and reactions to the future. *J. abnorm. soc. Psychol.* 1933–1934, **28**, 181–193.

ISRAELI, N. Measurement of the outlook on the future. *Psychol. Bull.*, 1934, **31**, 643–646

ISRAELI, N. Distress in the outlook of Lanca-

shire and Scottish unemployed. *J. appl. Psychol.*, 1935, **19**, 67–69. (a)

ISRAELI, N. *The outlook upon the future of British unemployed, mental patients and others.* Lancaster, England: Science Press, 1935. (b)

ISRAELI, M. *Abnormal personality and time.* New York: Science Press, 1936.

JOHNSTON, HARRIET M. A comparison of time estimation of schizophrenic patients with that of normal individuals. Unpublished masters dissertation, Brown Univer., 1939.

KAWASIMA, S. The influence of time intervals upon the perception of arm motion. *Jap. J. Psychol.*, 1937, **12**, 270–289. (Abstract)

KETCHUM, J. D. Time, values and social organization. *Canad. J. Psychol.*, 1951, **5**, 97–109.

KIERSEN, CELIA. Time and culture concepts in schizophrenic patients. *AMA Arch. Neurol. Psychiat.* 1951, **66**, 654–655. (Abstract)

KOWALSKI, W. J. The effect of delay upon the duplication of short temporal intervals. *J. exp. Psychol.*, 1943, **33**, 239–246.

LECOMTE, DU NOUY, P. *Biological time.* New York: Macmillan, 1937.

LECOMTE DU NUOY, P. Aging and physiological time. *Z. Altersforsch*, 1939, **1**, 301–309.

LECOMTE DU NUOY, P. Physiological time. *Proc. Amer. Phil. Soc.*, 1944, **87**, 435–437.

LESHAN, L. L. Time orientation and social class. *J. abnorm. soc. Psychol.*, 1952, **47**, 589–592.

LEVINE, M., & SPIVACK, G. Incentive, time conception and self-control in a group of emotionally disturbed male adolescents. *Amer. Psychologist*, 1957, **12**, 377.

LEWIN, K. Time perspective and morale. In G. Watson (Ed.), *Civilian morale.* Boston: Houghton Mifflin, 1942. Pp. 48–70.

LEWIN, K. Behavior and development as a function of the total situation. In L. Carmichael (Ed.), *Manual of child psychology.* New York: Wiley, 1946.

LEWIN, K. *Field theory and social science.* New York: Harper, 1951.

LEWIS, A. The experience of time in mental disorder. *Proc. Roy. Soc. Med.*, 1931–1932, **25**(1), 611–620.

LHAMON, W. T., & GOLDSTONE, S. The time sense. *AMA Arch. Neurol. Psychiat.*, 1956, **76**, 625–629.

LIPMAN, R. S. Some relationships between manifest anxiety, "defensiveness," and future time perspective. Paper read at East. Psychol., Assn., 1957.

LOEHLIN, J. C. The influence of different activities on the perception of time. Unpublished doctoral dissertation, Univer. of Calif., Berkeley, 1956.

LOOMIS, E. A. Space and time perception and distortion in hypnotic states. *J. Pers.*, 1951, **1**, 283–293.

MEERLOO, A. M. Father time. *Psychiat. Quart.* 1948, **22**, 587–608.

MEERLOO, A. M. Father time II. *Psychiat. Quart*, 1950, **24**, 657–671.

MINKOWSKI, E. Bergson's conceptions as applied to psychopathology. *J. nerv. ment. Dis.*, 1926, **63**, 553–568.

MOWRER, O. H., & ULLMAN, A. D. Time as a determinant in integrative learning. *Psychol. Rev.*, 1945, **52**, 61–90.

NAKAJIMA, S. On the time error in the successive comparison of time. *Jap. J. Psychol.* 1951, **21**, 36–45.

OAKDEN, E. C., & STUART, MARY. The development of the knowledge of time in children. *Brit. J. Psychol.* 1922, **12**, 309–336.

OBERNDORF, C. P. Time—its relation to reality and purpose. *Psychoanal. Rev.*, 1941, **28**, 139–155.

OLERON, G. Influence of the intensity of a sound on the estimation of duration. *Ann. psychol.*, 1952, **52**, 383–386.

PERLS, F. S. *Ego hunger and aggression.* London: Allen & Unwin, 1947.

PHILLIP, B. R. The anchoring of absolute judgments of short temporal intervals. *Bull. Canad. Psychol. Ass.*, 1944, **4**, 25–28.

PHILLIP, B. R. The effect of interpolated and extrapolated stimuli on the time order error in the comparison of temporal intervals. *J. genet. Psychol.*, 1947, **36**, 173–187.

PHILLIP, B. R., & LYTTLE, J. W. The relation between reaction time and duplication time for short intervals. *Bull. Canad. Psychol. Ass.*, 1945, **5**, 40–42.

PIGAET, J. The development of time concepts in the child. In P. H. Hoch & J. Zubin. (Eds.), *Psychopathology of childhood.* New York: Grune & Stratton, 1955.

PIÉRON, H. The sense of time and the chemical clock from the bee to man. *Livre jubilaire de M. L. Bouvier.* Paris: Firmin Didot, 1936.

PISTOR, F. Measuring time concepts of children. *J. educ. Res.*, 1939, **33**, 293–300.

PISTOR, F. How time concepts are acquired by children. *Educ. Method*, 1940, **20**, 107–112.

POSTMAN, L. Estimates of time during a series of tasks. *Amer. J. Psychol.*, 1944, **57**, 421–424.

POSTMAN, L., & MILLER, G. A. Anchoring of temporal judgments. *Amer. J. Psychol.*, 1945, **58**, 43–53.

PRATT, C. C. *The logic of modern psychology*, New York: MacMillan, 1948.

RABIN, A. I. Time estimation of schizophrenics and non-psychotics. *J. clin. Psychol.*, 1957, **13**, 88–90.

RAPAPORT, D. *Diagnostic psychological testing*. Vol. I. Chicago: Year Book, 1946.

RAPAPORT, D. *Emotions and memory*. New York: Int. Univer. Press, 1950.

RODGERS, W., & HAMMERSLEY, J. M. The consistency of stop-watch-study practitioners. *Occup. Psychol., Lond.*, 1954, **28**, 61–76.

ROELOFS, C. O., & ZEEMAN, W. P. The subjective duration of time intervals. *Acta psychol.*, 1949, **6**, 126–177.

ROELOFS, C. O., & ZEEMAN, W. P. Influential sequences of optical stimuli on the estimation of duration of a given interval of time. *Acta psychol.*, 1951, **8**, 89–128.

ROSS, S., & KATCHMAR, L. The construction of a magnitude function for short time intervals. *Amer. J. Psychol.*, 1951, **64**, 397–401.

SCHECHTER, D. E., SYMONDS, M., & BERNSTEIN, I. Development of the concept of time in children. *J. nerv. ment. Dis.*, 1955, **121**, 301–310.

SCHILDER, P. Psychopathology of time. *J. nerv. ment. Dis.*, 1936, **83**, 530–546.

SCHNEIDER, D. E. Time-space and the growth of the sense of reality. *Psychoanal. Rev.*, 1948, **35**(3), 229–251.

SCOTT, W. C. Some psychodynamic aspects of disturbed perception of time. *Brit. J. med. Psychol.*, 1948, **21**, 110–120.

SMITH, MARION W. Different cultural concepts of past, present and future. *Psychiatry*, 1952, **15**, 395–400.

SMYTHE, ELIZABETH J., & GOLDSTONE, S. The time sense: A normative, genetic study of the development of time perception. *Percept. mot. Skills*, 1957, **7**, 49–59.

SOLOMON, A. The relation of time estimation to personality traits. Unpublished master's thesis, Univer. of Kentucky, 1950.

SPRINGER, DORIS. Development in young children of an understanding of time and the clock. *J. genet. Psychol.*, 1952, **80**, 83–96.

STEINBERG, HANNAH. Changes in time perception induced by an anesthetic drug. *Brit. J. Psychol.*, 1955, **46**, 273–279.

STRAUSS, E. W. Disorders of personal time in depressive states. *South. Med. J.*, 1947, **40**(3), 154–158.

TEAHAN, J. E. Future time perspective, optimism and academic achievement. *J. abnorm. soc. Psychol.*, 1958, **57**, 379–380.

TURCHIOE, RITA M. The relation of adjacent inhibitory stimuli to the central tendency effect. *J. genet. Psychol.*, 1948, **39**, 3–14.

VAN DER HORST, V., Concerning the psychology of Korsakoff's syndrome. *Monatschr. Psychiat. Neurol.*, 1932, **83**, 65–84.

WALLACE, M. Future time perspective in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, **52**, 240–245.

WEINSTEIN, A. D., GOLDSTONE, S. & BOARDMAN, W. K. The effect of recent and remote frames of reference on temporal judgments of schizophrenic patients. *J. abnorm. soc. Psychol.*, 1958, **57**, 241–244.

WERNER, H. *Comparative psychology of mental development*. (Rev. ed.) Chicago: Follett, 1948.

WESLEY, E. B. *Teaching social studies*. Boston: Heath, 1942.

WOLFE, R. Space-and-time-experience under abnormal conditions, especially in mescal intoxication. *Dtsch. med. Wschr.*, 1952, **77** (6), 168–170.

WOLFF, W. *Values and personality*. New York: Grune & Stratton, 1950.

WOODROW, H. Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1224–1236.

WRIGHT, D. J., GOLDSTONE, S., BOARDMAN, W. K., & LHAMON, W. T. The time sense: Differential effect of series determined anchors on schizophrenics and normals. *Amer. Psychologist*, 1958, **13**, 357. (Abstract)

YATES, SYBILLE. Some aspects of time difficulties and their relation to music. *Int. J. Psycho-Anal.*, 1938, **16**, 341–354.

YOUNG, SHIRLEY A., & SUMNER, F. C. Personal equation, frame of reference, and other observations in remote reproduction of filled and unfilled time intervals. *J. genet. Psychol.*, 1954, **51**, 333–337.

(Received April 10, 1959)

PSYCHOLOGICAL BULLETIN
Vol. 57, No. 3, 1960

POSTASYMPTOTIC PERFORMANCE DECREMENTS DURING MASSED REINFORCEMENTS[1,2]

WILLIAM F. PROKASY
Pennsylvania State University

There are many secondary sources which, during discussions of massed reinforcement, refer to postasymptotic decrements in performance after extended training. One recent example of this is found in Deese (1958): "If, during conditioning, trials are massed very close together, the animals may slow their rates of responding and even stop altogether, even though the reinforcement is still present . . . " (p. 53). Similar statements are made by Bugelski (1956, pp. 166, 287, 349), Deese (1952, p. 56), Hilgard and Marquis (1940, p. 146), Hovland (1950, p. 616) and Hull (1943, p. 189f). The tenor of the discussions is that under conditions of massed reinforcement, sufficient numbers of trials will eventually lead to a decrement in the strength of a conditioned response. Some authors have labeled the genesis of this phenomenon "inhibition of reinforcement," a term coined by Hovland in his *Science* article of 1936.

Hull has treated "inhibition of reinforcement" much more formally in his *Principles of Behavior* (1943). His thirteenth corollary (1943, p. 189–191) is: "In the case of closely massed reinforcements, the curve of acquisition of effective excitatory potential . . . will be distorted by inhibition of reinforcement . . . in extreme cases showing an actual fall with continued practice." This theoretical position is in essential accord with the descriptions mentioned in the secondary sources cited above. In general, it would be predicted that continued, massed reinforcements will produce enough inhibition (built up with response evocation) to eventually offset the effects of the reinforcements, thus resulting in a gradual performance decrement even though reinforcement is not withdrawn. As a term, inhibition of reinforcement is not mentioned in Hull's final system (Hull, 1952); however, as Gleitman, Nachmias and Neisser (1954) point out, the Hullian construct of conditioned inhibition clearly implies that performance in a learning situation will asymptote, then gradually decline with continued, massed reinforcements. In fact, once performance has reached asymptote, the decline should occur equally rapidly with or without continued reinforcements.

We therefore have a situation where current secondary sources accept, as demonstrated, performance decrements during the massed reinforcement of a conditioned response and where a theoretical position predicts these decrements on the assumption of inhibition building up each time a response is made without adequate time for it to dissipate between trials. Deese (1958, p. 53) goes so far as to call this a "matter of experimental fact." It is the purpose of this paper to examine carefully the evidence for a performance decrement (as herein labeled for the rest of the paper) as a function of massed reinforcements. Classical and instrumental conditioning will be consid-

[1] This paper has been facilitated by a grant from the National Science Foundation.

[2] The author wishes to acknowledge the helpful comments provided by John F. Hall and William M. Lepley.

ered separately, to be followed by general comments and conclusions.

Classical Conditioning Studies

Historically, as might be expected, performance decrements were first observed by Pavlov (1927, p. 234f). Pavlov is cited in several secondary sources (Bugelski, 1956; Hovland, 1950; Hull, 1943) as providing data which indicate performance decrements with massed reinforcements, and it is this proposition that will be examined.

Several facts of the Pavlov data have much import for performance decrements and the conditions under which they might occur. First, conditioned responses learned to a conditioned stimulus of 30 seconds' duration very often spontaneously extinguish despite continued reinforcements. The characteristic of this response decrement and ultimate loss is its gradually increased latency until it is temporally coincident with the onset of the food (unconditioned stimulus). With further reinforcements, a test trial (conditioned stimulus alone) will not elicit the conditioned response even during the time the unconditioned stimulus would ordinarily have been presented. The intertrial interval in these situations is usually at 10 minutes or more on any given day of experimentation, and the performance decrement itself may require weeks, months or even years to show up.

Second, when a conditioned response is established to a conditioned stimulus of 10 seconds' duration, changing the duration of the conditioned stimulus to 30 seconds results in a loss of the conditioned response.

Third, a conditioned response learned to multiple conditioned stimuli may manifest a performance decrement to one of these stimuli, even to the extent that the unconditioned stimulus (food) is rejected, yet will occur in full strength to any one of the other stimuli to which the response has been learned. Further, when the former conditioned stimulus precedes one which still results in the response, a performance decrement occurs to the latter, also.

Fourth, there are two ways to reinstate a spontaneously extinguished conditioned response. If the response has been learned to a 30-second conditioned stimulus, it is possible to reduce the interstimulus interval to three or five seconds for several trials after which it is found that the response will follow the conditioned stimulus under the prior 30-second interstimulus interval. There are occasions, with highly trained animals, where all conditioned stimuli cease to have an effect and the animals even reject food which follows the conditioned stimulus presentation. That the animals are not satiated is indicated by the fact that they accept food presented in isolation from the conditioned stimulus. Under these circumstances a conditioned salivary response can be reinstated by employing entirely new conditioned stimuli: the response is quickly learned and stabilized to such new conditioned stimuli.

Finally, establishing a conditioned response to several reinforced stimuli, presented in order at 10-minute intervals, and then repeating one of those stimuli a number of times in a row while not employing the others may lead to a performance decrement to that stimulus (though it is still being reinforced), but leave the response to the other stimuli essentially unaffected. If this stimulus is repeatedly reinforced at intertrial intervals of one-and-a-half minutes, a response decrement will occur; but

presenting one of the other stimuli within that size interval will still result in the occurrence of a conditioned response. Further, when the 10-minute intertrial interval is restored, the responses again come back in full strength.

Several features of Pavlov's data prevent one from making the quick generalization that repeated, massed reinforcements in and of themselves will result in performance decrements. First the manner in which the conditioned response spontaneously extinguishes via increased latency, and the fact that conditioned responses do not occur when the conditioned stimulus is changed from a 10- to a 30-second duration suggest that temporal relationships between the conditioned and unconditioned stimuli may have an effect. It is well known, for instance, that large interstimulus intervals do not result in a well-established conditioned eyeblink response (e.g., Bernstein, 1934). There is the possibility that a response can be acquired to some degree of proficiency to a long duration conditioned stimulus, but that eventually the organism responds to the duration in the form of increased latency, finally resulting in a complete loss of the response because of an inability to discriminate such long durations accurately enough.

The facts that a conditioned response lost to one stimulus is unaffected in the presence of another stimulus and that a conditioned response lost to one stimulus can be trained to a new stimulus suggest more than the effects of continued reinforcement, massed or otherwise. The data indicate the possibility of an interaction between the reinforcement and a particular conditioned stimulus. In other words, if inhibition builds up, thus prohibiting the occurrence of the conditioned response, the inhibition is specific to a particular stimulus and is not solely the result of repeated massed reinforcements. It is to be noted that the conditioned stimuli employed by Pavlov were of varying effectiveness as far as degree of conditioning is concerned, and, further, that when performance decrements did occur depended to a certain extent upon the quality of the conditioned stimulus. A specific example of this (Pavlov, 1927, p. 235, 236) is when a response is established to a tactual stimulus within 179 trials. With a new conditioned stimulus (45-degree centigrade cutaneous stimulation) substituted, the conditioned response became established, but by the thirty-third trial the strength of the conditioned response had dropped to zero.

With respect to the massing factor, Pavlov generally employed an intertrial interval of 10 minutes or more between reinforcements. It is difficult to conceive of this as massed reinforcement.

The one feature of Pavlov's data which would appear to support the notion of performance decrements as a function of massed reinforcements is the result which obtains when the intertrial interval for an established conditioned response is changed from 10 to 1½ minutes. As pointed out above, this not only led to a performance decrement, but also to a rejection of food.[8] The occurrence of the response when another stimulus was introduced can be treated as an example of disinhibition. This experimental condition was not studied

[8] The fact that the organism could reject the unconditioned stimulus constitutes a departure from strict classical conditioning methodology. Just what this implies in the performance of highly trained *S*s, as were Pavlov's, is not known.

very often, as, according to Pavlov, recovery of the organism from the aftereffects of reinforcements was rapid (within one-and-a-half minutes) only for the particular dog yielding the results reported above. It is doubtful that the concept of performance decrement as a function of massed reinforcements should rest on the results of fewer than 30 massed trials on one animal; but, even in this animal, massing was not the sole factor operating as he rejected food on the last few trials with the original stimulus while accepting it when the disinhibiting stimulus was presented. In general, Pavlov's data provide a number of instances of performance decrements, but the degree to which massed reinforcements can be considered the major contributing factor is highly limited.

At least three sources (Bugelski, 1956; Deese, 1952, 1958) cite a study by Hovland (1936) as evidence for performance decrement.[4] Hovland administered 8 and 24 shock reinforcements to two groups of *Ss* in a classical GSR conditioning experiment. No acquisition data are presented, although when extinction followed the 24-trial acquisition series immediately there was an initial rise in the extinction curve. It is this rise that Hovland attributes to the aftereffects of inhibition of reinforcement. There are no data, however, in this study which indicate a performance decrement as herein defined. The only results given are extinction performances, where each curve is presented in terms of the first extinction response used as a base from which to compute the percentages of subsequent responses during the extinction trials. It would appear that, while Hovland did coin the phrase "inhibition of reinforcement," there is no evidence in his study for inhibition of reinforcement in its contemporary usage, contrary to the statements of the aforementioned authors.

Hilgard and Marquis (1940) cite Wendt (1930) as providing an example of performance decrement in classical conditioning. Wendt conducted an extensive investigation of the characteristics of the conditioned knee jerk, utilizing a blow to the right patellar tendon as a conditioned stimulus followed one-fifth of a second later by a blow to the left patellar tendon as an unconditioned stimulus. He discovered two kinds of response, distinguished primarily by latency. The first was a bilateral response with a latency of from 120 to 180 milliseconds. This particular response increased very rapidly in frequency over successive trials, then decreased during the final training trials. It would appear that it is to this response that Hilgard and Marquis refer. The second response took more trials to develop, and occurred with a latency of 200 to 500 milliseconds. This response was contralateral to the conditioned stimulus and ipsilateral to the unconditioned stimulus, as would be expected of a conditioned

[4] It is interesting to observe that the original definition of inhibition of reinforcement was not a performance decrement during massed reinforcements. Hovland (1936) states: "... two contrary varieties of curves of experimental extinction have been described. The first type is characterized by a continuous decline ... of the response on successive unreinforced elicitations.... The second type ... shows a larger response on the second or third extinction trial than on the first, demonstrating the initial rise characteristic of these curves.... There appears to be a negative adaptation to continuous reinforcement, an effect ... labeled ... 'inhibition of reinforcement'." In later work Hovland (1950, p. 616) discusses performance decrements, but does not use the label "inhibition of reinforcement" as do other authors (Bugelski, 1956; Deese, 1952, 1958; Razran, 1955; Solomon, 1948).

response, and did *not* decrease in frequency with repeated reinforcements. It is doubtful that the former response would be considered as a conditioned response by today's standards. Its latency and some of its trial-by-trial characteristics strongly resemble those of the "alpha" response (Grant, 1942, 1945; Prokasy & Truax, 1959), a reflex response to the conditioned stimulus in classical eyelid conditioning research, the frequency of which decreases with repeated stimulus presentations. The second response discovered by Wendt, however, not only did not drop out, but also falls within a latency range often delineated for the classically conditioned response (see, e.g., Prokasy, Grant, & Myers, 1958). Again, the evidence for a performance decrement as a function of massed reinforcements is slim.

Wolfle's now classic study of conditioned finger withdrawal (Wolfle, 1932) is cited by Hilgard (1933) as further evidence of this unique kind of extinction. Wolfle employed 72 *S*s spread over 13 interstimulus interval conditions. In four of these conditions the unconditioned stimulus preceded the conditioned stimulus, in eight the conditioned stimulus preceded the unconditioned stimulus and in the final condition the two stimuli were temporally coincident. Subjects were run in anywhere from two to nine experimental periods of 40 minutes' duration, the tests for conditioning being 10 irregularly spaced trials during the last 20 minutes of the session. The final results were that, of 72 Ss, 25 did not condition, 18 indicated no constant change from session to session, 8 showed a performance increment and 20 a performance decrement. Combining individual *S* data within each of the eight forward conditioning groups indicates that under several experimental conditions an over-all numerical decrease in percentage of conditioned responses occurred. In none of these conditions could the drop be considered significant with the possible exception of the .3-second interstimulus interval group. However, the facts that, (*a*) over one-third of the *S*s did not condition, (*b*) performance decrements were, at best, slight and, (*c*) the data were highly variable, all militate against accepting this study as serious evidence for a reliable performance decrement as a function of massed reinforcements.

Techniques of classical conditioning that are relatively standard today were employed by Hilgard (1933) who investigated the conditioned eyeblink response to a conditioned stimulus of light and an unconditioned stimulus of sound in two *S*s. One of the *S*s not only conditioned, but also manifested a decrement in amplitude of response after the 150th trial (50 trials a day for eight days) which was not regained. In addition to the fact that only one *S* gave this result (even though he constituted 50% of the sample), there is the distinct possibility that the unconditioned response to the tone as an unconditioned stimulus simply adapted out, the loss of the conditioned response being a function of the decreased motivational properties of the tone. It is already known that the reflex response to light adapts out (Grant, 1942, 1945; Prokasy & Truax, 1959), and it is not unlikely to have occurred to the tone, also. Be that as it may, the present study is difficult to generalize because of a one-*S* limitation.

The Newhall and Sears study (1933), cited as evidence for a performance decrement by Kantrow

(1937), is actually a psychophysical investigation comparing the results of a standard psychophysical technique for establishing a visual threshold with those of a conditioned finger withdrawal technique using weak visual stimuli as the conditioned stimuli. There were seven *Ss*: one undergraduate and six professional psychologists (including Newhall and Sears). First, visual thresholds were obtained by the method of constant stimuli, then *Ss* were required to report whether or not they saw a visual stimulus at the time they were being shocked following the visual stimuli. Interstimulus intervals were two seconds for three *Ss*, one second for two *Ss* and a half-second for the other two *Ss*. (Note, here, that the interstimulus interval was not conducive to good conditioning for five of the seven *Ss*.) The three *Ss* for whom the performance decrement was observed were the three who did not actually condition (made very few conditioned responses) in the first place. Limited *N*, inappropriate interstimulus intervals and the near or below threshold level of the conditioned stimuli place serious limitations on the degree to which any of these data can be generalized as evidence of a performance decrement.

Kasatkin and Levikova (1935), mentioned as support of performance decrements during reinforced conditioning trials by Kantrow (1937), investigated conditioned sucking responses in three human infants varying in age from 11 to 30 days at the beginning of the study. The general procedure was to present a tone two or three seconds before thrusting a lactating nipple (artificial) into the infant's mouth for a period of 15 to 30 seconds. When the sucking (conditioned) responses were firmly established, the conditioned stimulus was extended to a 10- to 15-second duration prior to the introduction of the unconditioned stimulus. The *Ss* received from 70 to 120 sessions over four or five months, with intertrial intervals of 10 to 60 seconds within each session. In two *Ss* the conditioned response dropped out after it had been established. Aside from the small *N*, two difficulties present themselves. First, the duration of the conditioned stimulus was changed in the middle of the experiment, and, second, the longer duration interstimulus intervals are, as mentioned above, as a rule not compatible with a high level of conditioning.

Kantrow (1937), cited by Hilgard and Marquis (1940), conducted a much more extensive study of conditioned sucking responses in the human infant. There were four stages to each conditioning cycle. First there was a control period (intertrial interval) of 25 to 75 seconds, followed by a buzzer as a conditioned stimulus of five seconds' duration. Then came the lactating nipple in conjunction with the buzzer for another 15 seconds, followed by the milk alone for from 15 to 120 seconds. This procedure was repeated for each of 16 infants (ages 44 to 117 days at the beginning of the experiment) during a single feeding until the infant either finished the bottle or refused to drink any more. These trials were administered in place of the regular four-hour feeding schedule. Five of the infants were given more than 16 conditioning sessions, and each of these five emitted fewer sucking responses during the final several sessions. This decrease in sucking behavior during the five-second duration conditioned stimulus was correlated with wide fluctuations in gross bodily activity. While the data from this experiment

indicate a performance decrement, it is not at all clear just how hunger, as the motivation, operates in this situation. By contemporary standards it is unusual classical conditioning methodology to infuse the unconditioned stimulus for such long periods of time and so to design the research that responding decreases to zero during the course of each training session. These two features of the present study should be explored more fully.

Hull (1943) cites an unpublished study by Calvin in which, under massed reinforcement training, there is approximately a 2 to 4% drop in frequency of conditioned eyeblink responses from the 21–30 trial block to the 41–50 trial block. While Hull puts enough credence in this result, along with the results of Pavlov (1927), to aid in the development of his thirteenth corollary, it is doubtful that such a performance decrease is reliable. Furthermore, recent research into the distribution of practice in eyeblink conditioning is not consistent with this finding (see e.g., Prokasy, Grant, & Myers, 1958).

Razran (1955) reports a preliminary study in which a group of *Ss* being tested on the first two extinction trials after 36 reinforcements are superior in performance to a group of *Ss* being so tested after 72 reinforcements. This is put forward as evidence for "decrement through overreinforcement," or, in the present context, performance decrement. The particular response under investigation was the magnitude of a human salivary response, classically conditioned to a 30-second duration series of flashing lights. The unconditioned stimulus in his experiment was a small pretzel. Two *Ss* served in a 12-trial-a-day condition, and two served in a 24-trial-a-day condition. All *Ss* served on four successive days, the first three of which were devoted to either 12 or 24 conditioning trials, the last day consisting of 18 extinction trials. It is to be noted that the test for amount of conditioning was made 24 hours after the last conditioning trials, and that this test indicated that the two *Ss* in the 36-trial group salivated significantly more (by a *t* test!) than did the two *Ss* in the 72-trial group. While no knowledge of the motivational properties of the 72nd pretzel are available, more recent evidence suggests that the decrement observed by Razran is very possibly not a "true" performance decrement. Prokasy (1958) administered 20, 40 and 60 classical eyelid conditioning trials to three groups of *Ss* respectively, and administered extinction trials 24 hours later. Although the 60-trial group was superior in performance to either the 20-trial or 40-trial groups, the loss in performance over the 24-hour interval was far greater for the 60-trial group (60% drop vs. 40% drop for the 20-trial group). This greater performance drop (as measured in the first five extinction trials) of the 60-trial group placed their performance below that of the 20-trial group during extinction. As Razran provides no data on the performance of the 72-trial group at the end of conditioning as compared to the 36-trial group, the critical comparison for a performance decrement is lacking; namely, performance itself being lower at one point in training than it was at some earlier point.

Although the data from classical conditioning studies which support the notion of a performance decrement as a function of massed reinforcements are severely limited, there is no lack of studies involving reasonably extensive training providing negative results. One recent study,

in particular, stands out. Ross (1959) administered 200 massed (average intertrial interval of 15 seconds) training trials to 20 *Ss* in a classical eyelid conditioning experiment. There was no evidence of a performance decrement despite the fact that the number of massed trials at one session is greater than the number used by Razran (1955), Calvin (in Hull, 1943) or Pavlov (1927).

Instrumental Conditioning Studies

The relevant research in instrumental conditioning in support of performance decrements as a function of massed reinforcements has more to recommend it than do the classical conditioning studies. Calvin, Clifford, Clifford, Bolden and Harvey (1956) report that 24 rats simply quit running down a straight 10-foot runway toward a goal box despite continued reinforcement. These investigators employed two levels of food deprivation (half of the animals received a total of 10 grams of food a day and the other half received a total of 12 grams of food a day) and two levels of distribution of practice (three minutes between trials for half of the animals and no delay between trials for the other half), with six animals in each of the resulting four experimental conditions. The criterion of extinction (during continued reinforcements) was two successive trials in which the animals refused to run within five minutes of being placed in the start box. Thirty trials a day were administered to all animals until this criterion was met. In brief, the criterion was met by all animals from 5 to 94 days of training. These investigators also found that the 10-gram group required a greater number of trials to reach the criterion than did the 12-gram group, and that there was a significantly greater running time for the massed groups. No other differences were statistically significant.

There are some difficulties in both the design and the results of this study which should be mentioned. First, reinforcement was doled out by permitting each animal to eat from a food dish, containing the full daily ration, for a period of 10 seconds. At the end of each day's trials the remaining food was placed in the cage with the animals. This type of feeding procedure is to be contrasted to the usual methods of maintaining body weight or a certain number of actual hours of food deprivation before the training trials.

Second, when an animal did not run within the five-minute criterion, he was placed in the goal box rather than removed from the apparatus. It is not unreasonable that this treatment had some effect on the runway behavior.

Third, the authors cite their data as support for conditioned inhibition, yet do not discuss the theoretical implications of the fact that the massed and distributed groups of *Ss* did not differ significantly in the number of trials required to reach criterion.

Finally, and in contrast to the data in classical conditioning, no mention is made of *gradually increasing* running times which would be expected from both the Hullian framework of conditioned inhibition as well as the usual conception of inhibition of reinforcement in secondary sources (e.g., Deese, 1952, 1958; Hilgard & Marquis, 1940).

Keehn and Sabbagh (1958) ran two groups of rats in a conditioned avoidance situation. One group of animals ($N=5$) received 100 trials of light-shock sequence. If they made a treadmill-running response the shock

was cut off, and if the running response came prior to the onset of the shock, no shock was received. The other group of animals ($N=5$) also received 100 such trials, but they were administered 10 a day for 10 days with a three-minute average intertrial interval within each day. It was found that the first group made more avoidance responses than did the second; and, although over-all conditioning was low, the second group manifested a decrease in avoidance behavior after the fifth session. In other words, the performance decrement came, not in the massed training group, but in the spaced training group. While this result is not easily interpreted, Keehn and Sabbagh suggest that the low conditioning level of the spaced group may be due to the infrequent trials presented each day, that there were insufficient conditions for performance to have occurred and remained stable. Despite the fact that Calvin (1958) persists in calling this particular result a case of conditioned inhibition, neither the concept of conditioned inhibition nor that of inhibition of reinforcement can handle the finding that the spaced group deteriorated in performance while the massed group did not.

A well-conceived study in this area has been reported by Kendrick (1958a). Ten rats were given 30 trials a day in a 12-foot runway under 20 hours of water deprivation with .25 cc. of water reinforcement. Each *S* was given daily trials until the extinction (performance decrement) criterion of refusal to run within five minutes on two consecutive trials during the first five daily trials for three consecutive days was met. Three hours after being run, *S*s were given a half-hour ad libitum drinking period. All 10 *S*s eventually stopped running, the over-all mean days to criterion for the 10 animals being 33.2 days. A retest 36 days after extinction yielded no evidence of spontaneous recovery. As in the Calvin et al. (1956) study, the behavior extinguished abruptly. A plot of mean running times over the six days prior to the three extinction criterion days indicates that the transition was not gradual as would be expected from a Hullian point of view (see Gleitman, Nachmias & Neisser, 1954). Log running times changed from an over-all average of one the day prior to the extinction series to two and a half on the first of the three criterion days.

Prior research (Seigel, 1947; Stellar & Hill, 1952) indicates that the total 30-trial reinforcement of 7.5 cc. employed in this study is only slightly greater than what a 24-hour deprived animal will drink in five minutes and far less than the 34 cc. ad libitum daily drinking found by Stellar and Hill with nondeprived animals. Thus, low drive level does not appear to be a factor in the findings. While Keehn (1959) suggests that unreinforced licks of the watchglass which provided water may be responsible for the results, this is highly speculative in light of the fact that the animals were water deprived, and extinguished within the first five trials on the criterion days. Kendrick's unusual findings differ from the usual form of extinction in that cessation was abrupt and there was no spontaneous recovery.

In a recent note, Kendrick (1958b) informally reports that employing a $\frac{1}{4}$ cc. water reinforcement for traversing a 10-foot runway and pressing a lever requiring 60 grams of pressure leads to performance decrements under massed training conditions. He states that the running response extinguishes, but that the bar pressing

itself does not extinguish. No data are presented, hence no assessment of this report can be made.

As in the case of classical conditioning, there is no lack of evidence for persistent responding over extensive spans of time when motivation is properly controlled. The reams of raw data presented by Ferster and Skinner (1957) are clearly adequate testimonial of perseverative behavior on the part of at least two animal species.

Concluding Comments

The hypothesis examined in this paper is essentially that massed reinforcements will result in a postasymptotic decrement in performance of a learned behavior. In classical conditioning there are good examples of performance decrements, primarily from Pavlov (1927), but these decrements in general are not a function of massed trials. Other factors, such as an interaction between particular conditioned stimuli and the conditioned and unconditioned responses, seem to be operating. The two instances of a performance decrement related to massing are both based upon one *S* (Pavlov 1927; Hilgard, 1933), and are, therefore, severely limited in generality.

Two instrumental studies indicate sharp performance decrements with massed reinforcements. In neither study, however, is there a control condition which permits the deduction that the decrements are a function of massing per se. In the first study (Calvin et al., 1956) there were both massed and distributed groups of animals, but these groups did not differ in the number of trials to reach the performance decrement criterion. While Kendrick (1958a) employed massed trials, he did not have a distributed trials control group. Thus, while the animals stopped running, with no manipulated independent variable correlated with this result, there is no way to make any decision concerning which factors in the experimental situation would consistently produce such a decrement.

In general, then, the hypothesis examined cannot be accepted: to date there is insufficient evidence to conclude that massed reinforcements in and of themselves will produce a performance decrement in the presence of repeated reinforcement. The present studies do, however, indicate several avenues of research available in attempting to isolate the variables correlated with the obtained performance decrements. In classical conditioning, multiple conditioned stimuli ought to be used in an attempt to determine why Pavlov's data indicate that some stimuli interact with the particular response system to provide performance decrements. Another complicating possibility are the long-term effects of particular conditioned stimuli. If there are short-term inhibition effects, Pavlov's data suggest, in addition, a second inhibition component related to highly overtrained *S*s that spontaneously extinguish after years of performance. Very little is known, furthermore, about the characteristics of the unconditioned response in classical conditioning. Any change in the unconditioned response, such as adaptation to the unconditioned stimulus or changes in the recovery time from the unconditioned response, could well influence conditioned response performance.

In instrumental conditioning, situations similar to that of Kendrick (1958a) ought to be more fully explored with variations in intertrial interval, length of runway, and amount of reinforcement. It remains to be seen, also, whether or not the sudden performance decrements ob-

served in the straight runway will occur in other instrumental response situations. If any general concept of a performance decrement is to be developed, clearly it must not be peculiar to the straight runway.

REFERENCES

BERNSTEIN, A. L. Temporal factors in the formation of conditioned eyelid reactions in human subjects. *J. gen. Psychol.*, 1934, **10**, 173–197.

BUGELSKI, B. R. *Psychology of learning.* New York: Holt, 1956.

CALVIN, A. D. Inhibition of reinforcement: Some comments. *Psychol. Rep.*, 1958, **4**, 677–678.

CALVIN, A. D., CLIFFORD, L. T., CLIFFORD, B., BOLDEN, L. & HARVEY, J. Experimental validation of conditioned inhibition. *Psychol. Rep.*, 1956, **2**, 51–56.

DEESE, J. *Psychology of learning.* New York: McGraw-Hill, 1952.

DEESE, J. *Psychology of learning.* (2nd ed.), New York: McGraw-Hill, 1958.

FERSTER, C., & SKINNER, B. F. *Schedules of reinforcement.* Appleton-Century-Crofts, 1957.

GLEITMAN, H., NACHMIAS, J., & NEISSER, U. The S-R reinforcement theory of extinction. *Psychol. Rev.*, 1954, **61**, 23–33.

GRANT, D. A. Sensitization and association in eyelid conditioning. *J. exp. Psychol.*, 1942, **32**, 201–212.

GRANT, D. A. A sensitized eyelid reaction to the conditioned eyelid response. *J. exp. Psychol.*, 1945, **35**, 393–402.

HILGARD, E. R. Modification of reflexes and conditioned reactions. *J. gen. Psychol.*, 1933, **9**, 210–215.

HILGARD, E. R., & MARQUIS, D. M. *Conditioning and learning.* Appleton-Century-Crofts, 1940.

HOVLAND, C. I. "Inhibition of reinforcement" and phenomena of experimental extinction. *Proc. Nat. Acad. Sci., Wash.*, 1936, **22**, 430–433.

HOVLAND, C. I. Human learning and retention. In Stevens, S. S. (Ed.), *Handbook of experimental psychology.* Wiley, 1950.

HULL, C. I. *Principles of behavior.* Appleton-Century-Crofts, 1943.

HULL, C. I. *A behavior system.* Yale Univer. Press, 1952.

KANTROW, R. W. An investigation of conditioned feeding responses and concomitant adaptive behavior in young infants. *Univer. Ia. Stud. Child Welf.*, 1937, **13**, No. 337.

KASATKIN, N. I., & LEVIKOVA, A. M. On the development of early conditioned reflexes and differentiations of auditory stimuli in infants. *J. exp. Psychol.*, 1935, **18**, 1–19.

KEEHN, J. D. More about inhibition of reinforcement. *Psychol. Rep.*, 1959, **5**, 141–142.

KEEHN, J. D., & SABBAGH, U. Conditioned inhibition and avoidance learning. *Psychol. Rep.*, 1958, **4**, 547–552.

KENDRICK, D. C. Inhibition with reinforcement (conditioned inhibition). *J. exp. Psychol.*, 1958, **56**, 313–318. (a)

KENDRICK, D. s_{IR} and drive level: a reply to Keehn and Sabbagh. *Psychol. Rep.*, 1958, **4**, 646. (b)

NEWHALL, S. M., & SEARS, R. R. Conditioning finger retraction to visual stimuli near the absolute threshold. *Comp. Psychol. Monog.*, 1933, **9**, No. 43.

PAVLOV, I. P. *Conditioned reflexes.* Oxford Univer. Press, 1927.

PROKASY, W. F. Extinction and spontaneous recovery of conditioned eyelid responses as a function of amount of acquisition and extinction training. *J. exp. Pscyhol.*, 1958, **56**, 319–324.

PROKASY, W. F., GRANT, D. A., & MYERS, N. A. Eyelid conditioning as a function of UCS intensity and intertrial interval. *J. exp. Psychol.*, 1958, **55**, 242–246.

PROKASY, W. F., & TRUAX, C. B. Reflex and conditioned responses as a function of manifest anxiety. *Amer. J. Psychol.*, 1959, in Press.

RAZRAN, G. Partial reinforcement of salivary CR's in adult human subjects: Preliminary study. *Psychol. Rep.*, 1955, **1**, 409–416.

ROSS, L. E. The decremental effects of partial reinforcement during acquisition of the conditioned eyelid response. *J. exp. Psychol.*, 1959, **57**, 74–82.

SEIGEL, P. S. The relationship between voluntary water intake, body weight loss and number of hours of water privation in the rat. *J. comp. Psychol.*, 1947, **40**, 231–238.

SOLOMON, R. L. The influence of work on behavior. *Psychol. Bull.*, 1948, **45**, 1–40.

STELLAR, E., & HILL, J. H. The rat's rate of drinking as a function of water deprivation. *J. comp. physiol. Psychol.*, 1952, **45**, 96–102.

WENDT, G. R. An analytical study of the conditioned knee-jerk. *Arch. Psychol.*, 1930, **19**, No. 123.

WOLFLE, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, **7**, 80–103.

(Received May 14, 1959)

ERRATUM

In the article "Partial Reinforcement: A Selective Review of the Literature Since 1950" by Donald J. Lewis (*Psychol. Bull.*, 1960, **57**, 1–28), the figure in the sixth line of page 19 should read 75%.

PSYCHOLOGICAL MONOGRAPHS

1959—Volume 73

MASS MEDIA AND CHILDREN: A Study of Exposure Habits and Cognitive Needs
by Lotte Bailyn. No. 471. $1.50

THE EFFECTS OF EARLY EXPERIENCE ON SUBSEQUENT EMOTIONALITY AND RESISTANCE TO STRESS
by Robert Ader. No. 472. $1.00

QUASI-THERAPEUTIC RELATIONS IN SMALL COLLEGE AND MILITARY GROUPS
by Fred E. Fiedler, Edwin B. Hutchins, and Joan S. Dodge. No. 473. $1.00

THE INFLUENCE OF DIFFERENT ACTIVITIES ON THE APPARENT LENGTH OF TIME
by John C. Loehlin. No. 474. $1.00

THE ONTOGENY OF CERTAIN LOGICAL OPERATIONS: Piaget's Formulation Examined by Nonverbal Methods
by Martin S. Braine. No. 475. $1.50

CONGRUENCIES AMONG INTERPRETATIONS OF PSYCHOLOGICAL TEST AND ANAMNESTIC DATA
by Kenneth B. Little and Edwin S. Shneidman. No. 476. $1.50

A Q-SORT STUDY OF THE VALIDITY OF EVALUATIONS MADE FROM PROJECTIVE TECHNIQUES
by Lloyd H. Silverman. No. 477. $1.00

THE INFLUENCE OF CLASSIFICATION REQUIREMENTS ON GRADIENTS OF RESPONSE
by Michael A. Wallach. No. 478. $1.00

A FACTOR-ANALYTIC STUDY ACROSS THE DOMAINS OF REASONING, CREATIVITY, AND EVALUATION
by Norman W. Kettner, J. P. Guilford, and Paul R. Christensen. No. 479. $1.00

GROUP PSYCHOTHERAPY AND CONCOMITANT PSYCHOLOGICAL CHANGE
by Earl J. Ends and Curtis W. Page. No. 480. $1.00

PHOTOSENSITIZATION IN RELATION TO MEAN AND STANDARD DEVIATION VALUES
by Richard T. Zegers. No. 481 $1.00

CORRELATES OF LATER PERFORMANCE AND SPECIALIZATION IN PSYCHOLOGY
by E. Lowell Kelly and Lewis R. Goldberg. No. 482.

Order from: AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street, N.W.
Washington 6, D.C.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN, U.S.A.